



TESIS - KI142502

**MODIFIKASI METODE EKSTRAKSI FITUR *MEL*
FREQUENCY CEPSTRAL COEFFICIENT UNTUK
IDENTIFIKASI PEMBICARA PADA LINGKUNGAN
BERDERAU MENGGUNAKAN RESIDU *ENDPOINT*
*DETECTION***

Yanuar Risah Prayogi
5113201031

DOSEN PEMBIMBING
Prof. Dr. Ir. Joko Lianto Buliali

PROGRAM MAGISTER
JURUSAN TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2015



THESIS - KI142502

MODIFIED FEATURE EXTRACTION METHOD MEL FREQUENCY CEPSTRAL COEFFICIENT FOR SPEAKER IDENTIFICATION ON NOISY ENVIRONMENT USING RESIDUE OF ENDPOINT DETECTION

Yanuar Risah Prayogi
5113201031

SUPERVISOR
Prof. Dr. Ir. Joko Lianto Buliali

MASTER PROGRAM
DEPARTMENT OF INFORMATICS
FACULTY OF INFORMATION TECHNOLOGY
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2015

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Komputer (M.Kom.)
di
Institut Teknologi Sepuluh Nopember Surabaya

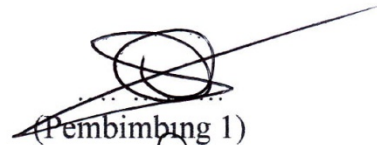
Oleh:
Yanuar Risah Prayogi
NRP. 5113201031

Dengan judul
**MODIFIKASI METODE EKSTRAKSI FITUR *MEL FREQUENCY*
CEPSTRAL COEFFICIENT UNTUK IDENTIFIKASI PEMBICARA PADA
LINGKUNGAN BERDERAU MENGGUNAKAN RESIDU *ENDPOINT*
*DETECTION***

Tanggal Ujian 5 Juni 2015
Periode Wisuda September 2015

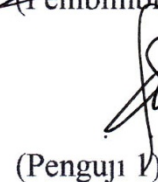
Disetujui oleh

Prof. Dr. Ir. Joko Lianto Buliali, M.Sc.
NIP. 196707271992031002



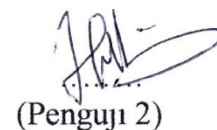
(Pembimbing 1)

Dr. Darlis Heru Murti, S.Kom, M.Kom.
NIP. 197712172003121001



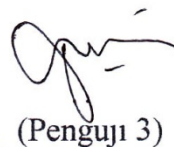
(Penguji 1)

Ahmad Saikhu, SSi, MT.
NIP. 197107182006041001



(Penguji 2)

Dini Adni Navastara, S.Kom, M.Sc.
NIP. 5100201405003



(Penguji 3)

Direktur Program Pasca Sarjana,

Prof. Dr. Ir. Adi Soeprijanto, MT.
NIP. 196404051990021001

MODIFIKASI METODE EKSTRAKSI FITUR *MEL FREQUENCY CEPSTRAL COEFFICIENT* UNTUK IDENTIFIKASI PEMBICARA PADA LINGKUNGAN BERDERAU MENGGUNAKAN RESIDU *ENDPOINT DETECTION*

Nama mahasiswa : Yanuar Risah Prayogi
NRP : 5113201031
Pembimbing : Prof. Dr. Ir. Joko Lianto Buliali, M.Sc

ABSTRAK

Salah satu permasalahan pada sistem identifikasi pembicara adalah fitur yang dihasilkan kurang tahan terhadap derau. Di lingkungan berderau, kinerja sistem identifikasi pembicara bisa turun secara signifikan. Hal ini disebabkan oleh perbedaan lingkungan ketika pelatihan dan pengujian. Salah satu metode ekstraksi fitur yang digunakan untuk identifikasi pembicara dan sensitif terhadap derau adalah *Mel Frequency Cepstral Coefficient* (MFCC). Di lingkungan bersih, kinerja yang dihasilkan oleh metode MFCC sangat tinggi, tetapi turun drastis ketika berada di lingkungan berderau.

Pada penelitian ini diusulkan memodifikasi metode MFCC menggunakan residu dari algoritma deteksi *endpoint*. Hasil dari algoritma deteksi *endpoint* adalah *speech* dan *nonspeech* (residu). *Nonspeech* atau residu ini biasanya tidak dipakai pada proses berikutnya. Pada sinyal suara yang berderau, residu dari algoritma deteksi *endpoint* sebagian besar diisi oleh derau itu sendiri sehingga bisa dijadikan informasi derau. Residu tersebut diekstrak untuk mendapatkan besaran (*magnitude*) frekuensi derau. Besaran frekuensi derau digunakan untuk menghilangkan derau pada sinyal utama atau *speech*.

Uji coba menggunakan 5 tipe derau dengan 7 tingkat SNR. Tipe derau yang digunakan adalah *f16*, *hfchannel*, *pink*, *volvo*, dan *white*. Sedangkan tingkat SNR yang digunakan adalah bersih, 25, 20, 15, 10, 5, dan 0 dB. Hasil uji coba menunjukkan bahwa metode yang diusulkan unggul pada mayoritas pembicara. Selain itu metode yang diusulkan juga unggul pada semua tipe derau dan unggul hampir pada semua tingkat SNR. Metode yang diusulkan menunjukkan rata-rata akurasi sebesar 14.69% lebih tinggi dari metode MFCC, 2.74% dari MFCC+Spectral Subtraction (SS), dan 6.4% dari MFCC+wiener.

Kata kunci : identifikasi pembicara, lingkungan berderau, *Mel Frequency Cepstral Coefficient* (MFCC), residu *endpoint detection*.

MODIFIED FEATURE EXTRACTION METHOD MEL FREQUENCY CEPSTRAL COEFFICIENT FOR SPEAKER IDENTIFICATION IN NOISES ENVIRONMENT USING RESIDUE OF ENDPOINT DETECTION

Name : Yanuar Risah Prayogi
Student Identity Number : 5113201031
Supervisor : Prof. Dr. Ir. Joko Lianto Buliali, M.Sc

ABSTRACT

One of the problems in the speaker identification system is a feature that generated less resistant to noise. In the noisy environment, the speaker identification system performance can drop significantly. It is caused by environmental differences when training and testing. One feature extraction method used to identify the speaker and sensitive to noise is Mel frequency cepstral coefficient (MFCC). In a clean environment, the performance generated by MFCC method is very high, but dropped dramatically when in the noisy environment.

In this study, we propose to modify the MFCC method using endpoint detection residues. Results of endpoint detection algorithm is speech and nonspeech (residue). Nonspeech or residues are usually not used in the next process. At the noisy signal, the residue of endpoint detection algorithm is filled by the noise itself so that it can be used as information noise. The residue is extracted to get the magnitude of the noisy signal. Magnitude of the noisy signal is used to remove noise on the main signal or speech.

The experiments using 5 types of noise with 7 levels of SNR. The type of noise that used is f16, hfchannel, pink, volvo, and white. While the level of SNR that used is clean, 25, 20, 15, 10, 5, and 0 dB. Experimental results show that the proposed method superior to the majority of the speakers. In addition the proposed method is also superior to all types of noise and superior in nearly all levels of SNR. The proposed method shows the average accuracy 14.69% higher than MFCC, 2.74% higher than MFCC+Spectral Subtraction (SS), and 6.4% higher than MFCC+wiener.

Keywords : speaker identification, noisy environment, mel frequency cepstral coefficient (MFCC), residue of endpoint detection.

KATA PENGANTAR

Alhamdulillah, segala puji bagi Allah SWT yang telah memberikan rahmat dan pertolongan-Nya sehingga penulis dapat menyelesaikan buku tesis ini dengan baik. Meski dalam menyelesaikan buku ini banyak ditemui kesulitan, namun berkat bimbingan dan bantuan berbagai pihak akhirnya penulis berhasil menyelesaikan buku ini. Pada kesempatan ini penulis ingin mengucapkan terima kasih kepada pihak-pihak yang membantu penulis dalam penulisan buku tesis ini sebagai berikut.

1. Orang tua penulis serta adik yang selalu senantiasa mendoakan dan dukungan untuk melanjutkan pendidikan yang lebih tinggi.
2. Prof. Dr. Ir. Joko Lianto Buliali, M.Sc selaku pembimbing pertama dan Dosen wali yang telah sabar membimbing penulis serta meluangkan waktu untuk memberikan saran berharga.
3. Waskitho Wibisono, S.Kom, M.Eng, Ph.D selaku Kepala Program Studi Pascasarjana Teknik Informatika ITS yang memberikan arahan dan masukan selama penulis menempuh pendidikan S2.
4. Dosen penguji, Dr. Darlis Heru Murti, S.Kom, M.Kom, Ahmad Saikhu, SSi, MT, dan Dini Adni Navastara, S.Kom, M.Sc yang telah memberikan masukan berharga dan kritikan yang membangun untuk tesis ini.
5. Kepada para Dosen pengajar, terima kasih atas segala ilmu yang telah diberikan kepada penulis selama penulis menempuh pendidikan S2.
6. Staf administrasi Program Pascasarjana Teknik Informatika, Bu Rini dan Bu Feni terima kasih atas informasi, pengertian dan kebijaksanaannya dalam urusan administrasi perkuliahan dan tesis.
7. Staf Laboratorium Pascasarjana Teknik Informatika, Pak Kunto atas bantuannya yang telah diberikan selama penulis melakukan uji coba di Laboratorium.
8. Teman-teman seperjuangan dan seangkatan (2013), Laili, Saiful, Pak Stendy, dan teman-teman lainnya yang tidak bisa disebutkan satu-persatu, terima kasih atas bantuan, perhatian, dan dukungan yang telah diberikan. Terima kasih juga karena telah bersedia menjadi teman diskusi yang baik sekaligus sebagai tempat untuk mencari motivasi sehingga penulis semangat mengerjakan tesis ini.

Akhirnya, Penulis berharap semoga penelitian ini bisa memberikan manfaat bagi berbagai pihak terutama bagi ilmu pengetahuan untuk pengembangan riset di bidang Komputasi Cerdas dan Visualisasi.

Surabaya, Maret 2015

Yanuar Risah Prayogi

DAFTAR ISI

ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvii
DAFTAR KODE SUMBER	xix
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Batasan Masalah	4
1.4 Tujuan dan Manfaat Penelitian	4
1.5 Kontribusi	4
BAB 2 KAJIAN PUSTAKA DAN DASAR TEORI	5
2.1 Sistem Identifikasi Pembicara	5
2.2 Deteksi <i>Endpoint</i>	6
2.3 <i>Mel Frequency Cepstral Coefficient</i> (MFCC)	7
2.3.1 <i>Preemphasis</i>	8
2.3.2 <i>Frame Blocking</i>	9
2.3.3 <i>Windowing</i>	9
2.3.4 <i>Fast Fourier Transform</i> (FFT)	10
2.3.5 <i>Mel-frequency Filtering</i>	10
2.3.6 <i>Discrete Cosine Transform</i> (DCT)	12
2.3.7 <i>Delta & Delta-delta Cepstrum</i>	12
2.3.8 <i>Log Energy</i>	13
2.4 <i>Gaussian Mixture Model</i> (GMM)	13
2.4.1 Estimasi Parameter	14
2.4.2 Identifikasi Pembicara	15
2.5 <i>Signal to Noise Ratio</i> (SNR)	15
BAB 3 METODE PENELITIAN	17
3.1 Rancangan Penelitian	17
3.2 Perancangan Sistem	18
3.2.1 <i>Praproses</i> (Deteksi <i>Endpoint</i>)	19
3.2.2 Modifikasi Metode MFCC	19
3.2.2.1 <i>Preemphasis</i>	20
3.2.2.2 <i>Frame Blocking</i>	21
3.2.2.3 <i>Windowing</i>	21
3.2.2.4 <i>Fast Fourier Transform</i> (FFT)	21
3.2.2.5 Rata-rata <i>Magnitude</i> Frekuensi	22
3.2.2.6 Pengurangan <i>Magnitude</i> Frekuensi	22
3.2.2.7 <i>Mel Frequency Filtering</i>	22
3.2.2.8 <i>Discrete Cosine Transform</i> (DCT)	23
3.2.2.9 <i>Delta & Delta-delta Cepstrum</i>	23
3.2.3 Pelatihan dan Pengenalan	23

3.3 Rancangan Uji Coba	23
3.3.1 Data Uji Coba.....	24
3.3.2 Pengujian dan Evaluasi	25
BAB 4 HASIL DAN PEMBAHASAN	29
4.1 Pengolahan <i>Dataset</i>	29
4.2 Implementasi Sistem Identifikasi Pembicara.....	31
4.2.1 Spesifikasi Sistem	31
4.2.2 Deteksi <i>Endpoint</i>	31
4.2.3 <i>Preemphasis</i>	32
4.2.4 <i>Frame Blocking</i>	33
4.2.5 <i>Windowing</i>	33
4.2.6 <i>Fast Fourier Transform</i> (FFT).....	34
4.2.7 Rata-rata <i>Magnitude</i> Frekuensi	34
4.2.8 Pengurangan <i>Magnitude</i> Frekuensi.....	34
4.2.9 <i>Mel-Frequency Filtering</i>	35
4.2.10 <i>Discrete Cosine Transform</i> (DCT).....	35
4.2.11 <i>Delta & Delta-delta Cepstrum</i>	35
4.3 Uji Coba Skenario Identifikasi Pembicara.....	36
4.3.1 Nilai Variabel/Parameter pada Uji Coba.....	37
4.3.2 Hasil Uji Coba Skenario Identifikasi Pembicara	39
4.3.3 Analisis Hasil Uji Coba Skenario Identifikasi Pembicara	40
4.3.4 Uji <i>t</i> Hasil Analisis Sistem Identifikasi Pembicara	47
BAB 5 KESIMPULAN DAN SARAN	51
5.1 Kesimpulan	51
5.2 Saran	52
DAFTAR PUSTAKA.....	53
LAMPIRAN 3	79
BIODATA PENULIS	91

DAFTAR TABEL

Tabel 3.1 Nilai Variabel pada Algoritma Deteksi <i>Endpoint</i>	20
Tabel 4.1 Nilai Variabel/Parameter pada Uji Coba	38
Tabel 4.2 Nilai Variabel/Parameter pada Algoritma <i>Endpoint Detection</i>	38
Tabel 4.3 Nilai Uji t Akurasi Metode Ekstraksi Fitur Terhadap Pembicara (t kritis = 1.69)	39
Tabel 4.4 Nilai Uji t Akurasi Metode Ekstraksi Fitur Terhadap Tipe Derau (t kritis = 2.13)	39
Tabel 4.5 Nilai Uji t Akurasi Metode Ekstraksi Fitur Terhadap Tingkat SNR (t kritis = 1.94)	40
Tabel 4.6 Nilai Standar Deviasi Akurasi Metode Ekstraksi Fitur Terhadap Pembicara	42
Tabel 4.7 Akurasi Metode Ekstraksi Fitur Terhadap Tipe Derau (dalam satuan %)	43
Tabel 4.8 Nilai Standar Deviasi dan Rata-rata Akurasi Metode Ekstraksi Fitur Terhadap Tipe Derau	43
Tabel 4.9 Akurasi Metode Ekstraksi Fitur Terhadap Tingkat SNR (dalam satuan %)	45

DAFTAR KODE SUMBER

Kode Sumber 4.1 Deteksi <i>Endpoint</i>	32
Kode Sumber 4.2 <i>Preemphasis</i>	33
Kode Sumber 4.3 <i>Frame Blocking</i>	33
Kode Sumber 4.4 <i>Windowing</i>	33
Kode Sumber 4.5 <i>Fast Fourier Transform</i> (FFT).....	34
Kode Sumber 4.6 Rata-rata <i>Magnitude</i>	34
Kode Sumber 4.7 Pengurangan <i>Magnitude</i>	35
Kode Sumber 4.8 <i>Mel-Frequency Filtering</i>	35
Kode Sumber 4.9 <i>Discrete Cosine Transform</i> (DCT)	36
Kode Sumber 4.10 <i>Delta & Delta-delta Cepstrum</i>	36

DAFTAR GAMBAR

Gambar 2.1 Contoh Sistem Identifikasi Pembicara	6
Gambar 2.2 Contoh Hasil Deteksi <i>Endpoint</i>	7
Gambar 2.3 Metode MFCC (Young, 1996; Han dkk, 2006)	8
Gambar 2.4 Ilustrasi <i>Frame Blocking</i>	9
Gambar 2.5 <i>Gaussian Mixture Model</i> dengan M Komponen (Reynolds dan Rose, 1995)	14
Gambar 3.1 Arsitektur Sistem Identifikasi Pembicara.....	19
Gambar 3.2 Blok Proses pada Bagian Modifikasi Metode MFCC.....	20
Gambar 3.3 Pelatihan dan Pengenalan Menggunakan GMM.....	24
Gambar 3.4 Skenario Uji Coba ketika Pelatihan dan Pengenalan	26
Gambar 4.1 Grafik Akurasi Metode Ekstraksi Fitur Terhadap Tipe Derau.....	44
Gambar 4.2 Grafik Akurasi Metode Ekstraksi Fitur Terhadap Tingka SNR.....	46

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Identifikasi pembicara adalah proses mengenali identitas pembicara menggunakan suara yang diberikan (Leon, 2009). Pada sistem identifikasi pembicara, identitas pembicara sudah didaftarkan dalam *database*. Pembicara mengklaim sebuah identitas menggunakan suaranya. Sedangkan sistem mencocokkan suara pembicara dengan *database* yang ada pada sistem. Identifikasi pembicara ini bisa digunakan sebagai *password* pada sistem keamanan. Misalkan pada sistem keamanan perangkat *smartphone*, untuk membuka dan menggunakan perangkat tersebut diperlukan identifikasi pengguna. Sehingga hanya pengguna tertentu saja yang mempunyai hak dalam penggunaan perangkat tersebut.

Kelemahan utama pada sistem identifikasi pembicara adalah kurang tahan terhadap lingkungan yang berderau (Zhang dan Abdulla, 2007). Derau adalah suara yang tidak diinginkan. Di dunia nyata, kinerja sistem identifikasi pembicara bisa turun secara signifikan. Hal ini disebabkan oleh perbedaan lingkungan ketika pelatihan dan pengujian. Ketika pengujian, derau pada sinyal suara ternyata lebih besar atau lebih kecil dibandingkan ketika pelatihan sehingga kinerja dari sistem identifikasi pembicara menurun.

Mel Frequency Cepstral Coefficient atau MFCC adalah metode yang terkenal dan umum digunakan untuk ekstraksi fitur suara (Han dkk, 2006). *Cepstral coefficient* yang dihasilkan sensitif terhadap derau (Zhang dan Abdulla, 2007). Pada lingkungan bersih, kinerja yang dihasilkan oleh metode MFCC sangat tinggi, tetapi turun drastis ketika berada di lingkungan berderau. Oleh karena itu dibutuhkan modifikasi pada metode MFCC.

Modifikasi MFCC pernah dilakukan oleh Mark D. Skowronski dan John G. Harris (Skowronski dan Harris, 2002) yang memodifikasi pada lebar pita (*bandwidth*) dari *Mel-scale filter bank*. Lebar pita *Mel-scale filter bank* yang awalnya sesuai dengan *Mel-scale* diperlebar sehingga lebih banyak yang tumpang

tindih (*overlap*). Wang Yutai dkk (Yutai dkk, 2009) melakukan modifikasi dengan menambahkan *pitch* sebagai fitur yang dinamik. Tomi Kinnunen (Kinnunen dkk, 2012) memodifikasi pada proses *windowing*. Proses *windowing* menggunakan beberapa *window (taper)* di domain waktu untuk menghasilkan beberapa spektrum di domain frekuensi. Kizuki Matsumoto dkk (Matsumoto dkk, 2014) memodifikasi dengan membagi lebar pita (*bandwidth*) frekuensi sinyal suara menjadi dua yaitu rendah dan tinggi menggunakan analisis frekuensi modulasi. Garis besar dari penelitian tersebut adalah memodifikasi MFCC agar terbentuk fitur yang handal di lingkungan berderau. Rata-rata akurasi yang dihasilkan lebih tinggi dari pada metode MFCC. Meskipun demikian, akurasinya turun drastis seiring turunnya tingkat SNR.

Beberapa peneliti melakukan pendekatan lain dengan menambahkan *Spectral Subtraction* (SS) untuk menghilangkan sinyal derau. Shang-Ming Lee dkk (Lee dkk, 2002) selain menerapkan *Principal Component Analysis* (PCA) pada *Mel-scale filter bank* juga menambahkan *Spectral Subtraction* (SS) untuk mengurangi derau. Sedangkan Wu Zunjing dan Cao Zhigang (Zunjing dan Zhigang, 2005) mengganti fungsi log pada MFCC dengan fungsi pangkat. Selain itu juga menambahkan *Spectral Subtraction* (SS) pada praproses. Dengan penambahan *Spectral Subtraction*, akurasi dua penelitian diatas meningkat dibandingkan dengan tanpa penambahan *Spectral Subtraction*. Kelebihan *Spectral Subtraction* adalah dapat menghilangkan sinyal derau dengan memanfaatkan informasi sinyal derau yang diperoleh dari beberapa *frame* awal. Bagaimanapun akurasi yang dihasilkan tetap turun drastis seiring turunnya tingkat SNR. Selain itu jumlah *frame* awal yang digunakan harus tepat karena jika jumlah *frame* awal terlalu banyak maka sinyal suara ikut hilang.

Peneliti lainnya yang mencoba menghilangkan derau pada MFCC adalah Paresh M. Chauhan dan Nikita P. Desai (Chauhan dan Desai, 2014). Paresh M. Chauhan dan Nikita P. Desai menambahkan *Wiener filter* setelah proses FFT. *Wiener filter* bekerja di domain frekuensi karena terletak setelah proses FFT. Paresh M. Chauhan dan Nikita P. Desai menguji *Wiener filter* dengan meletakkan di domain waktu (setelah *framing*) dan di domain frekuensi (setelah FFT). Hasil yang diperoleh adalah peletakan *Wiener filter* di domain frekuensi (setelah FFT)

lebih efektif dibandingkan di domain waktu (setelah *framing*). Akurasi yang dihasilkan lebih tinggi dibandingkan dengan MFCC. Meskipun demikian, *Wiener filter* menghilangkan derau berdasarkan estimasi sehingga akurasi yang dihasilkan turun seiring turunnya tingkat SNR.

Dari beberapa penelitian sebelumnya, masih ada kebutuhan untuk meningkatkan kinerja metode MFCC baik pada lingkungan bersih maupun berderau. Untuk meningkatkan kinerja MFCC pada sistem identifikasi pembicara, baik pada tingkat SNR rendah maupun tinggi, diperlukan modifikasi yang mampu menghilangkan frekuensi derau sesuai dengan tingkat SNR. Maka diusulkan modifikasi MFCC menggunakan residu dari algoritma deteksi *endpoint*. Motivasi menggunakan residu dari algoritma deteksi *endpoint* adalah karena algoritma deteksi *endpoint* digunakan untuk memisahkan *speech* dengan *nonspeech* (Yong dan Leimin, 2011). *Nonspeech* ini biasanya dibuang dan tidak akan dipakai pada proses berikutnya sehingga bisa disebut sebagai residu. Di dalam sinyal yang berderau, residu dari algoritma deteksi *endpoint* sebagian besar diisi oleh derau (Wu dan Wang, 2005). Residu ini dapat dijadikan sebagai informasi sinyal derau seperti yang dilakukan oleh *Spectral Subtraction*. Residu algoritma deteksi *endpoint* diubah ke domain frekuensi sehingga didapatkan *magnitude* frekuensi derau. *Magnitude* tersebut kemudian digunakan untuk menghilangkan derau pada *speech* di domain frekuensi seperti yang dilakukan pada penelitian Paresh M. Chauhan dan Nikita P. Desai (Chauhan dan Desai, 2014).

1.2 Perumusan Masalah

Perumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana cara mendapat *magnitude* frekuensi derau dari residu algoritma deteksi *endpoint*?
2. Bagaimana cara memodifikasi metode MFCC agar dapat menghilangkan derau pada sinyal suara menggunakan residu algoritma deteksi *endpoint*?
3. Bagaimana pengaruh modifikasi metode MFCC terhadap kinerja sistem dalam hal akurasi?

1.3 Batasan Masalah

1. Sistem identifikasi pembicara dibangun menggunakan bahasa pemrograman MATLAB versi R2012a.
2. Teks yang digunakan ketika pelatihan dan pengenalan adalah sama (*text-dependent*).
3. Pembicara yang diidentifikasi adalah pembicara yang sudah terdaftar di database (*close-set*).
4. Data pelatihan dan pengenalan berupa file wav dengan frekuensi sampling 16000 Hz.
5. Diasumsikan setiap sinyal suara terdapat *nonspeech* diawal, diakhir, atau keduanya.
6. Derau yang digunakan adalah derau yang bertipe stasioner yaitu *white noise*, *pink noise*, *cockpit noise 3 (F-16)*, *HF channel noise*, *vehicle interior noise (volvo 340)*.
7. Tingkat SNR yang digunakan dalam penambahan derau adalah bersih, 25, 20, 15, 10, 5, dan 0 dB.
8. Dalam satu file uji hanya ditambahi satu tipe dengan satu tingkat SNR.

1.4 Tujuan dan Manfaat Penelitian

Tujuan dari penelitian ini adalah untuk meningkatkan kinerja dari metode ekstraksi fitur MFCC dalam hal akurasi, baik pada lingkungan yang bersih maupun berderau menggunakan residu dari algoritma deteksi *endpoint*.

Manfaat dari penelitian ini adalah agar sistem dapat mengidentifikasi pembicara, baik pada lingkungan bersih maupun berderau dengan tingkat akurasi yang tinggi.

1.5 Kontribusi

Kontribusi pada penelitian ini adalah memodifikasi metode ekstraksi fitur, *Mel Frequency Cepstral Coefficient* (MFCC), untuk identifikasi pembicara pada lingkungan yang bersih dan berderau menggunakan residu dari algoritma deteksi *endpoint*.

BAB 2

KAJIAN PUSTAKA DAN DASAR TEORI

Dalam bab ini akan diuraikan tentang dasar teori penelitian yang meliputi sistem identifikasi pembicara, praproses sinyal suara, metode ekstraksi fitur, pelatihan dan pengenalan. Praproses sinyal suara terdiri dari deteksi *endpoint*. Ekstraksi fitur menggunakan metode *Mel Frequency Cepstral Coefficient* (MFCC). Pada metode MFCC terdapat blok proses yaitu *frame blocking*, *preemphasis*, *windowing*, *fast fourier transform* (FFT), *mel-frequency filtering*, *discrete cosine transform* (DCT), *log energy*, dan *delta & delta-delta cepstrum*. Sedangkan pada pelatihan dan pengenalan digunakan metode *Gaussian Mixture Model* (GMM).

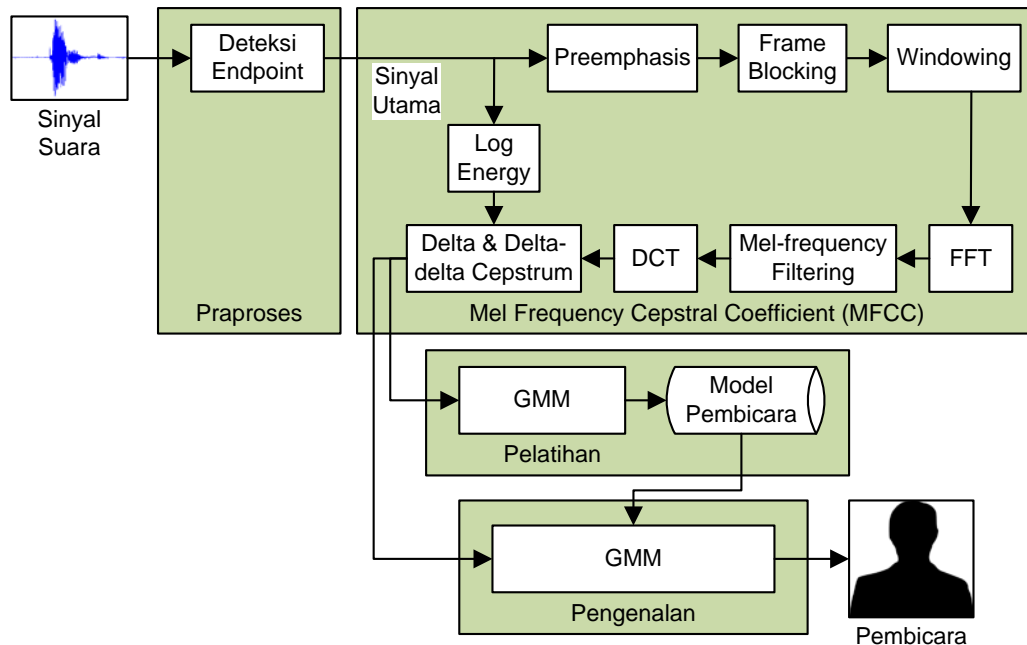
2.1 Sistem Identifikasi Pembicara

Identifikasi pembicara adalah proses mengenali pembicara yang sudah terdaftar di database menggunakan suara yang diucapkan. Identifikasi pembicara dibagi menjadi dua yaitu *text-dependent* dan *text-independent*. Pada sistem *text-dependent*, teks yang digunakan ketika pelatihan dan pengenalan adalah teks yang sama. Sebaliknya sistem *text-independent*, teks yang digunakan ketika pelatihan dan pengenalan bisa berbeda (Leon, 2009).

Sistem identifikasi pembicara secara umum dibagi menjadi 4 bagian atau blok yaitu praproses, ekstraksi fitur, pelatihan, dan pengenalan. Setiap bagian terdapat blok proses yang lebih kecil lagi. Misalkan didalam praproses terdapat deteksi *endpoint*. Pada ekstraksi fitur terdapat proses *preemphasis*, *frame blcoking*, *windowing*, *fast fourier transform* (FFT), *mel-frequency filtering*, *discrete cosine transform* (DCT), *log energy*, dan *delta & delta-delta cepstrum*. Sedangkan pada pelatihan dan pengenalan terdapat proses *Gaussian Mixture Model* (GMM).

Masukkan sistem berupa sinyal suara sedangkan keluaran berupa identitas pembicara. Hasil dari pelatihan berupa database model pembicara. Kemudian database model pembicara digunakan pada pengenalan sebagai

referensi untuk mencari kedekatan yang paling maksimum. Contoh sistem identifikasi pembicara ditunjukkan pada Gambar 2.1.



Gambar 2.1 Contoh Sistem Identifikasi Pembicara

2.2 Deteksi *Endpoint*

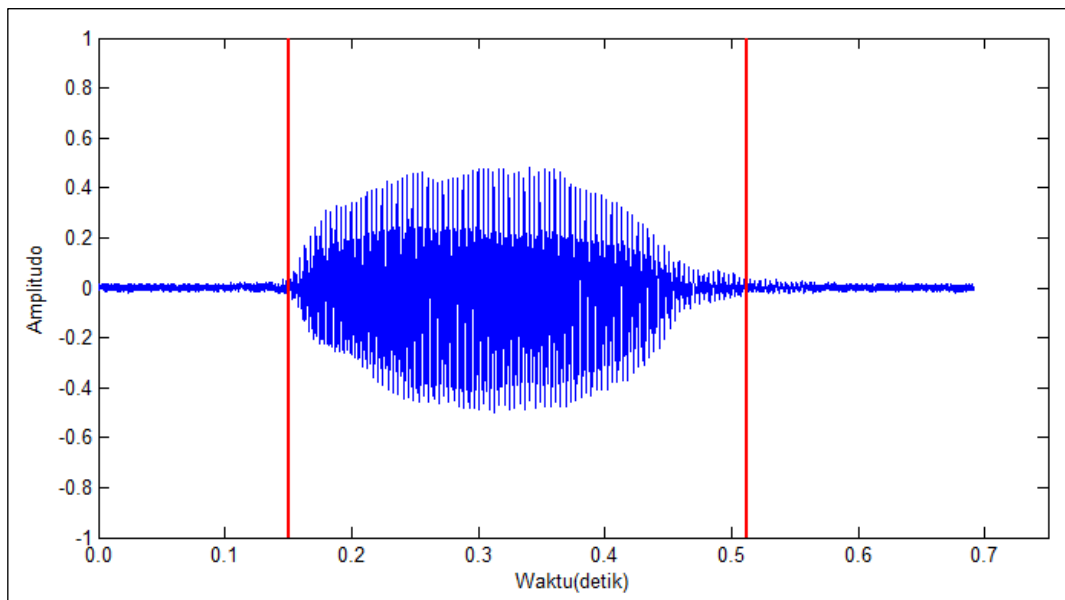
Praproses adalah proses yang dikerjakan pada tahap awal sebelum ekstraksi fitur. Praproses sinyal suara digunakan untuk meningkatkan kualitas sinyal suara. Meningkatkan sinyal suara dalam hal ini adalah menghilangkan bagian yang tidak diinginkan misalnya sinyal *nonspeech* atau sinyal derau. Praproses sinyal suara merupakan langkah penting untuk membentuk sistem pengenalan suara atau identifikasi pembicara yang handal dan efisien. Salah satu metode yang termasuk praproses adalah deteksi *endpoint*.

Deteksi *endpoint* adalah algoritma yang digunakan untuk memisahkan antara bagian *speech* dan *nonspeech* di sinyal suara (Yong dan Leimin, 2011). Fungsi dari algoritma ini adalah untuk mengambil bagian yang diinginkan yaitu bagian *speech*. Bagian *speech* adalah bagian yang bersih sinyal suara. Sedangkan *nonspeech* adalah daerah sunyi/hening. Pada lingkungan berderau, bagian *nonspeech* biasanya diisi oleh derau. Deteksi *endpoint* yang digunakan pada

penelitian ini berdasarkan *short-time energy*. *Short-time energy* menggambarkan kuantitas sinyal suara yang dinyatakan dengan kuadrat dari amplitudo. *Short-time energy* dinyatakan dengan persamaan (Yong dan Leimin, 2011) (2.1)

$$E_k = \sum_{m=0}^{N-1} x_k^2(m) = \sum_{m=0}^{N-1} [w(m)x(k+m)]^2, \quad 0 \leq m \leq N-1, \quad (2.1)$$

dimana E_k adalah energi pada frame ke- k , N adalah jumlah sampel dalam satu *frame*, $w(m)$ adalah fungsi window, dan $x_k(m)$ adalah sinyal suara. Contoh hasil dari deteksi *endpoint* ditunjukkan pada Gambar 2.2.



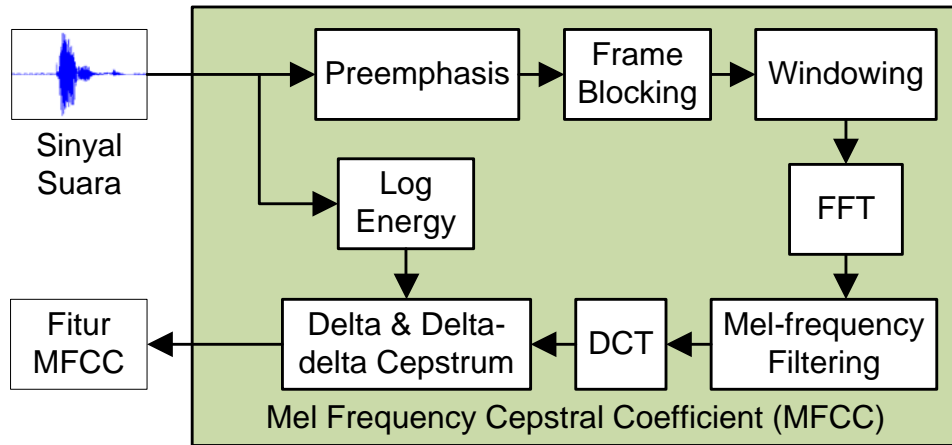
Gambar 2.2 Contoh Hasil Deteksi *Endpoint*

Garis merah pada Gambar 2.2 adalah pembatas antara *speech* dan *nonspeech*. *Speech* terletak dibagian tengah sedangkan *nonspeech* terletak dibagian awal dan akhir.

2.3 Mel Frequency Cepstral Coefficient (MFCC)

MFCC adalah metode ekstraksi fitur yang terkenal dan umum digunakan (Han dkk, 2006). Proses-proses yang ada pada metode MFCC adalah *preemphasis*, *frame blocking*, *windowing*, *Fast Fourier Transform* (FFT), *Mel-frequency Filtering*, *Discrete Continuous Transform* (DCT), *log energy*, dan *Delta*

& *Delta-delta Cepstrum* (Young, 1996; Han dkk, 2006). Proses-proses yang ada pada metode MFCC ditunjukkan pada Gambar 2.3.



Gambar 2.3 Metode MFCC (Young, 1996; Han dkk, 2006)

Masukkan dari metode ini berupa sinyal suara sedangkan keluaran berupa fitur MFCC. Satu *frame* menghasilkan satu vektor fitur sehingga satu sinyal suara akan menghasilkan beberapa baris vektor fitur.

2.3.1 *Preemphasis*

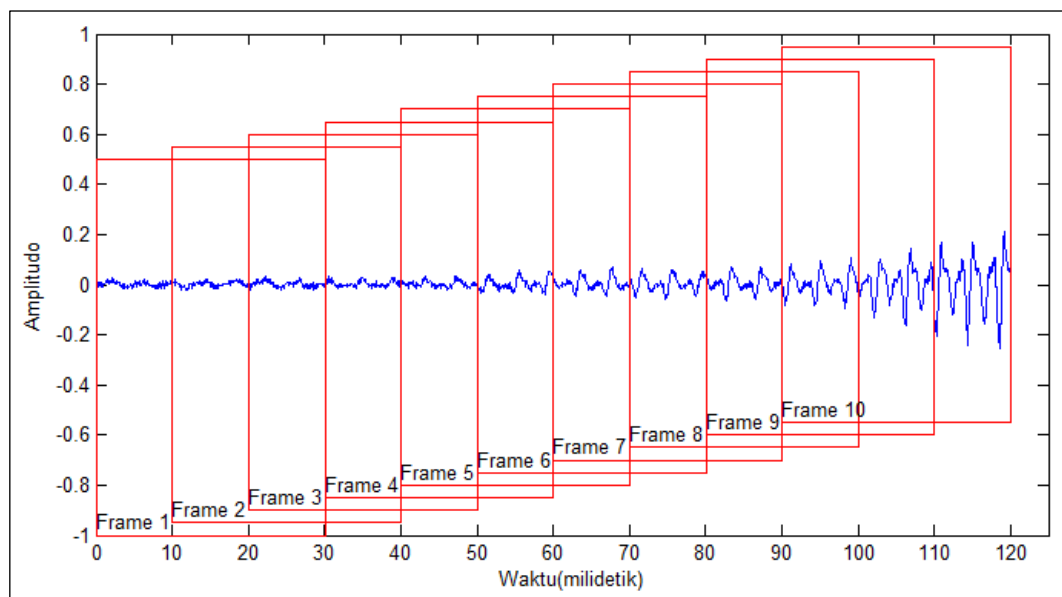
Pada sinyal suara yang asli biasanya memiliki energi frekuensi rendah yang cukup banyak. Sedangkan energi frekuensi tinggi sedikit. Perbedaan energi ini terjadi karena sinyal frekuensi rendah disampling dengan frekuensi sampling yang cukup tinggi sehingga menghasilkan nilai numerik yang sama. Tujuan dari *preemphasis* adalah meningkatkan energi dari frekuensi tinggi. Fungsi dari *preemphasis* hampir sama dengan high-pass filter yaitu untuk melewati komponen frekuensi tinggi dari sinyal suara. Misalkan $s(n)$ adalah sinyal suara maka, *preemphasis* dapat dinyatakan dengan persamaan (Vergin, 1995) (2.2)

$$y(n) = s(n) - \alpha * s(n - 1) , \quad (2.2)$$

dimana $y(n)$ adalah sinyal *preemphasis* dan α adalah konstanta yang bernilai 0.9 sampai 1. Nilai α yang sering digunakan adalah 0.97 (Han dkk, 2006).

2.3.2 Frame Blocking

Sinyal suara dibagi menjadi blok-blok kecil yang disebut *frame*. Antara *frame* satu dengan yang lain terjadi tumpang tindih (*overlapping*). Tiap *frame* terdiri dari N sampel dan tumpang tindih sebanyak M sampel. Pada metode MFCC panjang *frame* yang biasa digunakan adalah 30 milidetik atau $N = 480$ sampel untuk frekuensi sampling 16000 Hz. Sedangkan panjang tumpang tindih adalah 10 milidetik atau $M = 160$ sampel untuk frekuensi sampling 16000 Hz. Tujuan *frame blocking* adalah membagi sinyal suara menjadi *frame* kecil dengan sampel yang cukup untuk mendapatkan informasi yang cukup (Gupta, 2013). Jika ukuran *frame* terlalu kecil, informasi yang didapat tidak cukup untuk dipercaya. Sebaliknya jika terlalu besar, informasi didalam *frame* sering berubah-ubah (Gupta, 2013). Ilustrasi *frame blocking* sinyal suara ditunjukkan pada Gambar 2.4.



Gambar 2.4 Ilustrasi *Frame Blocking*

2.3.3 Windowing

Proses *windowing* dilakukan untuk tiap *frame*. *Windowing* digunakan untuk meminimalkan ketidaksinambungan pada awal dan akhir *frame* (Leon, 2009). Dengan menggunakan *window*, sinyal suara pada awal dan akhir *frame*

menjadi runcing (Gupta, 2013). Prosesnya adalah sinyal suara dikalikan dengan *window*. Proses perhitungan *windowing* ditunjukkan pada persamaan (2.3)

$$y(m) = w(m)x(m), \quad 0 \leq m \leq N - 1, \quad (2.3)$$

dimana $w(m)$ adalah fungsi *window* dan $x(m)$ adalah sinyal suara. Pada dasarnya banyak fungsi *window* yang bisa digunakan misalkan *blackman window*, *gaussian window*, dan *hamming window*. Pada metode MFCC, *window* yang sering digunakan adalah *hamming window* (Gupta, 2013). Fungsi *hamming window* dinyatakan dengan persamaan (Gupta, 2013) (2.4)

$$w(m) = 0.54 - 0.46 * \cos\left(\frac{2\pi m}{N-1}\right), \quad 0 \leq m \leq N - 1, \quad (2.4)$$

2.3.4 Fast Fourier Transform (FFT)

Fast Fourier Transform (FFT) merubah sinyal suara dari domain waktu ke domain frekuensi. FFT adalah algoritma tercepat yang mengimplementasikan *Discete Fourier Transform* (DFT). DFT dinyatakan dengan persamaan (Gupta, 2013) (2.5)

$$X(k) = \sum_{m=0}^{N-1} x(m)e^{-j2\pi km/N}, \quad k = 0, 1, 2, \dots, N - 1, \quad (2.5)$$

dimana j adalah bilangan imajiner yang bernilai $\sqrt{-1}$ dan $x(m)$ adalah sinyal suara. Frekuensi yang dihasilkan bergantung dari frekuensi sampling. Jika frekuensi sampling 16000 Hz maka frekuensi yang dihasilkan antara 0 sampai 8000 Hz. Hasil dari FFT adalah bilangan kompleks yang terdiri dari bilangan riil (a) dan imajiner (b). Sama dengan vektor, nilai besaran frekuensi (*magnitude*) didapatkan dengan $(\sqrt{a^2 + b^2})$ akar kuadrat a ditambah b .

2.3.5 Mel-frequency Filtering

Mel-frequency filtering adalah analisis frekuensi menggunakan *filter bank*. Hasil dari proses ini adalah frekuensi yang dipetakan menggunakan *Mel-filter*, $H(k, m)$. *Mel-frequency filtering* dinyatakan dengan persamaan (Junjin dan Junjun, 2011) (2.6).

$$D(m) = \sum_{k=0}^{N-1} |X(k)| * H(k, m), \quad m = 1, 2, \dots, M \quad (2.6)$$

dimana M adalah jumlah *filter bank*, N adalah panjang FFT, $X(k)$ adalah hasil FFT ke- k , dan $H(k, M)$ adalah *Mel-filter*. Nilai $H(k, m)$ dinyatakan dengan persamaan (Junqin dan Junjun, 2011) (2.7).

$$\begin{aligned} H(k, m) &= 0, \quad k < f[m - 1] \\ H(k, m) &= \frac{2(k - f[m - 1])}{(f[m + 1] - f[m - 1])(f[m] - f[m - 1])}, \\ &\quad f[m - 1] \leq k < f[m] \\ H(k, m) &= \frac{2(f[m + 1] - k)}{(f[m + 1] - f[m - 1])(f[m + 1] - f[m])}, \\ &\quad f[m] \leq k < f[m + 1] \\ H(k, m) &= 0, \quad k \geq f[m + 1] \end{aligned} \quad (2.7)$$

dimana $f[m]$ respon dari *Mel* frekuensi didapatkan dengan persamaan (Junqin dan Junjun, 2011) (2.8).

$$f[m] = \frac{N}{f_s} \text{Mel}^{-1} \left(\text{Mel}(f_l) + m \frac{\text{Mel}(f_h) - \text{Mel}(f_l)}{M + 1} \right), \quad (2.8)$$

dimana f_l dan f_h adalah frekuensi batas bawah dan frekuensi batas atas yang ada didalam *filter bank*, f_s adalah frekuensi sampling, $\text{Mel}(f)$ adalah *Mel-scale* dan $\text{Mel}^{-1}(f)$ adalah *invers Mel-scale*. *Mel-scale* dan *invers Mel-scale* didefinisikan dengan persamaan (Junqin dan Junjun, 2011) (2.9) dan (2.10).

$$\text{Mel}(f) = 2595 * \log_{10} \left(1 + \frac{f}{700} \right), \quad (2.9)$$

$$\text{Mel}^{-1}(f) = 700 * (10^{\frac{f}{2595}} - 1), \quad (2.10)$$

dimana f adalah frekuensi yang akan diskala menggunakan *Mel-scale*. Parameter f dalam satuan Hz.

2.3.6 Discrete Cosine Transform (DCT)

Setelah didapatkan spektrum Mel, dilakukan operasi DCT. DCT adalah proses yang digunakan untuk mengembalikan dari domain frekuensi ke domain waktu (Leon, 2009). DCT selain untuk mengembalikan ke domain waktu juga untuk kompresi spektrum (Gupta, 2013). Operasi DCT dapat dinyatakan dengan persamaan (Leon, 2009) (2.11)

$$C(k) = w(k) \sum_{m=1}^L D(m) \cos\left(\frac{\pi(2m-1)(k-1)}{2L}\right), k = 1, 2, \dots, N \quad (2.11)$$

dimana L adalah jumlah *Mel* spektrum *filter*, $D(m)$ adalah mel spektrum, dan N adalah jumlah koefisien *cepstrum*. Sedangkan $w(k)$ dinyatakan dengan persamaan 2.12. Untuk komponen pertama, $C(0)$, tidak dipakai karena mengandung sedikit informasi (Leon, 2009).

$$w(k) = \begin{cases} \frac{1}{\sqrt{L}}, & k = 1 \\ \sqrt{\frac{2}{L}}, & 2 \leq k \leq L \end{cases} \quad (2.12)$$

2.3.7 Delta & Delta-delta Cepstrum

Pendengaran manusia lebih sensitif terhadap karakteristik dinamik sinyal suara (Junqin dan Junjun, 2011). Turunan pertama (*delta*) dan turunan kedua (*delta-delta*) menggambarkan karakteristik dinamik dari sinyal suara. Nilai *delta cepstrum* didapatkan dari turunan pertama koefisien *cepstrum* yang dinyatakan dengan persamaan (Hutter, 1996) (2.13)

$$d(k) = \frac{\sum_{t=1}^T t * (C(k+t) - C(k-t))}{2 \sum_{t=1}^T t^2}, \quad 1 \leq k \leq N \quad (2.13)$$

dimana $C(k)$ adalah koefisien *cepstrum*, N adalah jumlah koefisien *cepstrum*, dan T adalah konstan. Nilai T yang biasa digunakan adalah 2 (Junqin dan Junjun, 2011). Sedangkan *delta-delta cepstrum* didapatkan dari turunan pertama *delta cepstrum*.

2.3.8 Log Energy

Log *energy* adalah operasi sederhana yang didapatkan dari log penjumlahan kuadrat amplitudo. Log *energy* (E) dapat dinyatakan dengan persamaan (Han, 2006) (2.14)

$$E = \log \sum_{n=1}^N s(n)^2 \quad (2.14)$$

dimana $s(n)$ adalah sinyal suara didalam *frame* dan N adalah panjang *frame*. Sinyal suara yang digunakan pada log *energy* adalah sinyal suara yang belum dilakukan operasi *preemphasis*. Kemudian dilakukan operasi *frame blocking* untuk membagi sinyal suara menjadi *frame-frame* kecil.

2.4 Gaussian Mixture Model (GMM)

Gaussian Mixture Model adalah model statistik yang umum digunakan pada pengenalan pembicara dan pengenalan suara. *Gaussian mixture density* didapatkan dari penjumlahan bobot sebanyak M komponen seperti yang ditunjukkan pada Gambar 2.5. Nilai *Gaussian Mixture* dinyatakan dengan persamaan (Reynolds dan Rose, 1995) (2.15).

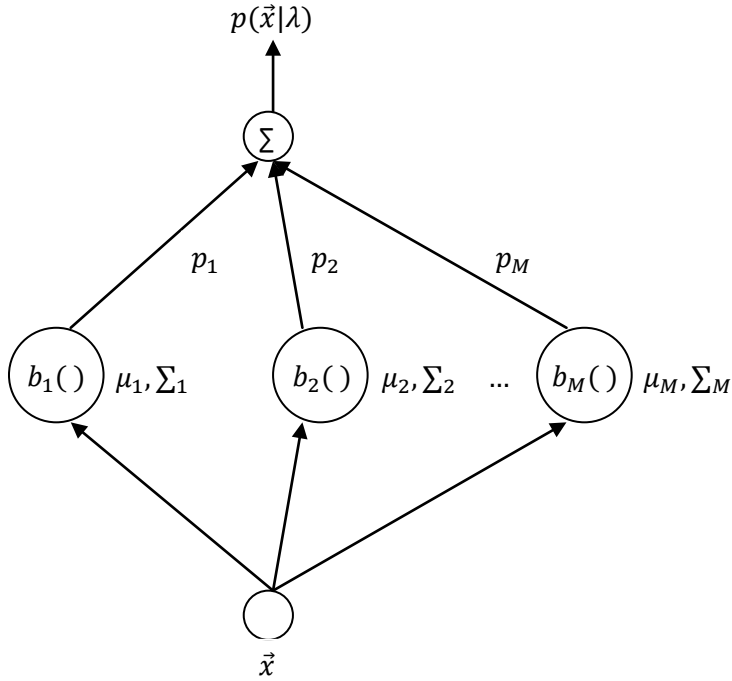
$$p(\vec{x}|\lambda) = \sum_{i=1}^M p_i b_i(\vec{x}), \quad (2.15)$$

dimana \vec{x} adalah vektor dengan D dimensi, $b_i(\vec{x})$ adalah komponen densitas, dan p_i adalah bobot *mixture* dimana memenuhi batasan $\sum_{i=1}^M p_i = 1$. Variabel $b_i(\vec{x})$ diperoleh dari persamaan (Reynolds dan Rose, 1995) (2.16).

$$b_i(\vec{x}) = \frac{1}{(2\pi)^{D/2} |\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{\mu}_i)^T \Sigma_i^{-1} (\vec{x} - \vec{\mu}_i) \right\} \quad (2.16)$$

dimana $\vec{\mu}_i$ adalah vektor rata-rata dan Σ_i adalah matrik kovarian. Parameter keseluruhan yang dipakai oleh Gaussian Mixture Model adalah bobot *mixture*, vektor rata-rata, dan matrik kovarian yang semuanya dinyatakan dengan persamaan (Reynolds dan Rose, 1995) (2.17).

$$\lambda = \{p_i, \vec{\mu}_i, \Sigma_i\}, \quad i = 1, \dots, M \quad (2.17)$$



Gambar 2.5 *Gaussian Mixture Model* dengan M Komponen (Reynolds dan Rose, 1995)

2.4.1 Estimasi Parameter

Tujuan dari pelatihan adalah membentuk model GMM dengan nilai parameter λ yang optimum. Untuk membentuk λ yang optimum digunakan algoritma *Expectation-Maximization* (EM). EM dimulai dengan inisialisasi awal λ dan estimasi nilai $\bar{\lambda}$ yang baru dimana $p(X|\bar{\lambda}) \geq P(X|\lambda)$. Nilai $\bar{\lambda}$ yang baru menjadi nilai inisialisasi awal untuk iterasi berikutnya dan prosesnya diulangi terus sampai konvergen atau mencapai batas *threshold* tertentu. Misalkan ada vektor data latih $X = \{\vec{x}_1, \dots, \vec{x}_T\}$ sebanyak T maka nilai $\bar{\lambda} = \{\bar{p}_i, \vec{\bar{\mu}}_i, \bar{\sigma}_i^2\}$ didapatkan dengan persamaan (Reynolds dan Rose, 1995) 2.18, 2.19, dan 2.20.

$$\bar{p}_i = \frac{1}{T} \sum_{t=1}^T p(i|\vec{x}_t, \lambda), \quad (2.18)$$

$$\bar{\mu}_i = \frac{\sum_{t=1}^T p(i|\vec{x}_t, \lambda) \vec{x}_t}{\sum_{t=1}^T p(i|\vec{x}_t, \lambda)}, \quad (2.19)$$

$$\bar{\sigma}_i^2 = \frac{\sum_{t=1}^T p(i|\vec{x}_t, \lambda) x_t^2}{\sum_{t=1}^T p(i|\vec{x}_t, \lambda)} - \bar{\mu}_i^2, \quad (2.20)$$

dimana σ_i^2 , x_t , dan μ_i secara berturut-turut adalah elemen dari vektor $\vec{\sigma}_i^2$, \vec{x}_t , dan $\vec{\mu}_i$. Sedangkan peluang posteriori didapatkan dengan persamaan (Reynolds dan Rose, 1995) 2.21.

$$p(i|\vec{x}_t, \lambda) = \frac{p_i b_i(\vec{x}_t)}{\sum_{k=1}^M p_k b_k(\vec{x}_t)}, \quad (2.21)$$

2.4.2 Identifikasi Pembicara

Misalkan ada kelompok identifikasi pembicara $S = \{1, 2, \dots, S\}$ yang direpresentasikan dengan parameter GMM $\lambda_1, \lambda_2, \dots, \lambda_S$. Tujuan dari identifikasi pembicara adalah mencari model GMM yang memiliki nilai peluang posteriori maksimum untuk rentetan observasi \vec{x} yang diberikan. Untuk mencari nilai peluang posteriori maksimum menggunakan persamaan (Reynolds dan Rose, 1995) 2.22.

$$\hat{S} = \arg \max_{1 \leq k \leq S} \sum_{t=1}^T \log p(\vec{x}_t | \lambda_k), \quad (2.22)$$

dimana $p(\vec{x}_t | \lambda_k)$ dijelaskan pada persamaan 2.15.

2.5 Signal to Noise Ratio (SNR)

Signal to Noise Ratio (SNR) adalah ukuran yang digunakan untuk membandingkan tingkat sinyal yang diinginkan dengan tingkat kebisingan latar belakang (*noise*). SNR didefinisikan dengan rasio antara power sinyal yang diinginkan dengan power sinyal latar belakang (*noise*) yang ditunjukkan pada persamaan 2.23.

$$SNR = \frac{P_{signal}}{P_{noise}} = \left(\frac{A_{signal}}{A_{noise}}\right)^2, \quad (2.23)$$

dimana P adalah power dan A adalah *root mean square* (RMS) amplitudo. Karena rentang nilai sinyal sangat lebar sehingga dituliskan dengan skala logaritma decibel (dB) yang ditunjukkan pada persamaan 2.24.

$$\begin{aligned} SNR_{dB} &= 10 \log_{10} \left(\frac{P_{signal}}{P_{noise}} \right) = 10 \log_{10} \left[\left(\frac{A_{signal}}{A_{noise}} \right)^2 \right] \\ &= 20 \log_{10} \left(\frac{A_{signal}}{A_{noise}} \right), \end{aligned} \quad (2.24)$$

BAB 3

METODE PENELITIAN

Dalam bab ini akan diuraikan tentang rancangan penelitian, rancangan sistem, dan rancangan uji coba.

3.1 Rancangan Penelitian

Secara umum, penelitian ini dilakukan dalam beberapa tahap yaitu diawali dari studi literatur, perumusan masalah dan hipotesis, perancangan metode dan implementasi, serta uji coba dan evaluasi. Sedangkan penulisan laporan penelitian dimulai dari awal sampai akhir penelitian ini.

1. Studi literatur

Mempelajari berbagai literatur tentang sistem identifikasi pembicara mulai dari metode praproses, metode ekstraksi fitur, dan metode pelatihan dan pengenalan.

2. Perumusan masalah dan hipotesis

Merumuskan permasalahan yang akan diteliti dan mencari solusinya. Mengungkapkan hipotesis mengenai solusi permasalahan yang diteliti.

3. Perancangan metode dan implementasi

Merancang dan mengimplementasikan metode penyelesaian dari permasalahan yang telah dirumuskan berdasarkan pengetahuan yang diperoleh dari studi literatur. Rancangan metode penyelesaian yang diusulkan akan dijelaskan pada sub bab 3.2.

4. Uji coba dan evaluasi

Melakukan pengujian dan evaluasi terhadap metode yang telah dirancang dengan menerapkan beberapa skenario. Uji coba dan evaluasi akan dijelaskan pada sub bab 3.3.

5. Penyusunan laporan

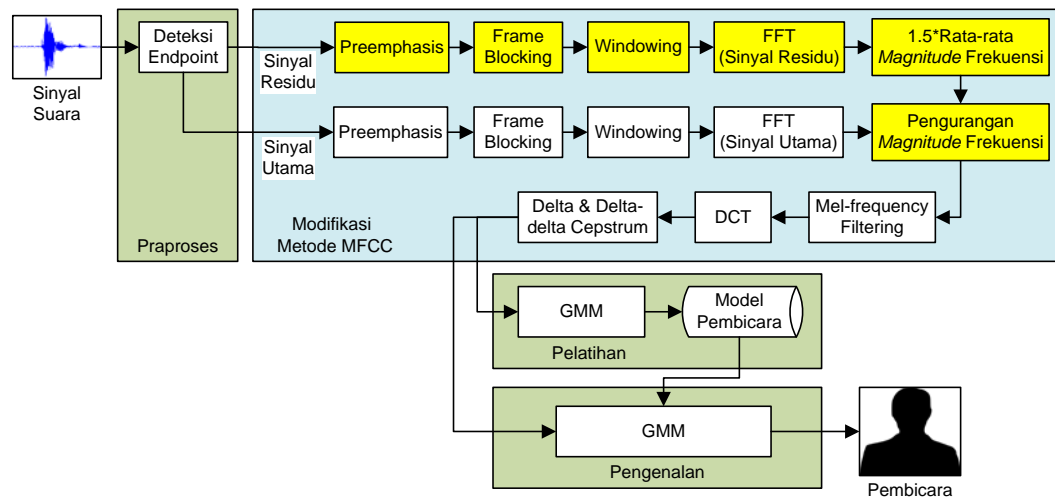
Penyusunan laporan dilakukan mulai dari awal sampai akhir penelitian ini. Penyusunan laporan ditulis dalam bentuk laporan tesis berdasarkan ketentuan yang berlaku.

3.2 Perancangan Sistem

Sistem identifikasi pembicara secara umum dibagi menjadi 4 bagian yaitu praproses, ekstraksi fitur, pelatihan, dan pengenalan. Arsitektur sistem identifikasi pembicara ditunjukkan oleh Gambar 3.1. Pada Gambar 3.1, bagian praproses, pelatihan, dan pengenalan yang ditandai warna hijau muda adalah bagian yang tidak mengalami modifikasi. Sedangkan bagian metode MFCC yang ditandai warna biru muda adalah bagian yang mengalami modifikasi. Bagian praproses terdiri dari blok proses deteksi *endpoint*. Sedangkan bagian pelatihan dan pengenalan menggunakan *Gaussian Mixture Model* (GMM). Proses *log energy* pada MFCC dihilangkan kemudian ditambahkan beberapa proses yang ditunjukkan dengan blok warna kuning.

Dasar pemikiran penambahan blok proses pada MFCC adalah metode *Spectral Subtraction* (SS) (Lim dan Oppenheim, 1979) dan penelitian Paresh dan Nikita (Chauhan dan Desai, 2014). Metode SS menghilangkan derau berdasarkan informasi yang didapatkan dari beberapa *frame* awal. Sedangkan Paresh dan Nikita menghilangkan derau menggunakan *Wiener filter* (Chauhan dan Desai, 2014). Hasil yang didapatkan adalah letak *Wiener filter* lebih efektif berada di domain frekuensi (setelah FFT) daripada di domain waktu (setelah *framing*).

Penambahan beberapa blok proses dimaksudkan untuk mengurangi derau di domain frekuensi. Blok proses *preemphasis*, *frame blocking*, *windowing*, dan FFT (sinyal residu) digunakan untuk mengekstrak frekuensi derau. Cara kerjanya sama dengan metode *Spectral Subtraction* yaitu mendapatkan informasi frekuensi derau dari beberapa *frame* awal. Bedanya, informasi frekuensi derau didapatkan dari residu deteksi *endpoint*. Kemudian dihitung rata-rata magnitude frekuensi untuk semua *frame*. Hasilnya digunakan untuk mengurangi magnitude frekuensi sinyal utama. Pengurangan ini terjadi di domain frekuensi seperti penelitian Paresh dan Nikita yang menghilangkan derau menggunakan *Wiener filter* di domain frekuensi (Chauhan dan Desai, 2014). Sedangkan blok proses *log energy* dihilangkan karena nilai *log energy* berubah-ubah tergantung tingkat sinyal derau. *Log energy* didapatkan dari penjumlahan kuadrat amplitudo sinyal suara. Semakin tinggi tingkat sinyal derau menyebabkan nilai *log energy* juga semakin tinggi.



Gambar 3.1 Arsitektur Sistem Identifikasi Pembicara

3.2.1 Praproses (Deteksi *Endpoint*)

Praproses adalah proses yang dikerjakan pada tahap awal sebelum ekstraksi fitur. Praproses sinyal suara digunakan untuk meningkatkan kualitas sinyal suara. Praproses yang digunakan pada penelitian ini adalah deteksi *endpoint*. Deteksi *endpoint* adalah algoritma yang memisahkan antara *speech* dan *nonspeech*. *Speech* berisi sinyal utama sedangkan *nonspeech* adalah daerah sunyi/hening. Daerah *nonspeech* ini bisa disebut juga sebagai residu dari algoritma deteksi *endpoint*.

Deteksi *endpoint* yang digunakan pada penelitian ini berdasarkan *short-time energy*. *Short-time energy* menggambarkan kuantitas sinyal suara yang dinyatakan dengan kuadrat dari amplitudo (Yong dan Leimin, 2011). Pada *short-time energy* terdapat variabel α , jumlah sampel dalam satu *frame* atau N , jumlah sampel yang tumpang tindih atau M , fungsi window atau W dan *threshold*. Nilai variabel pada algoritma deteksi *endpoint* yang digunakan pada penelitian ini ditunjukkan pada Tabel 3.1.

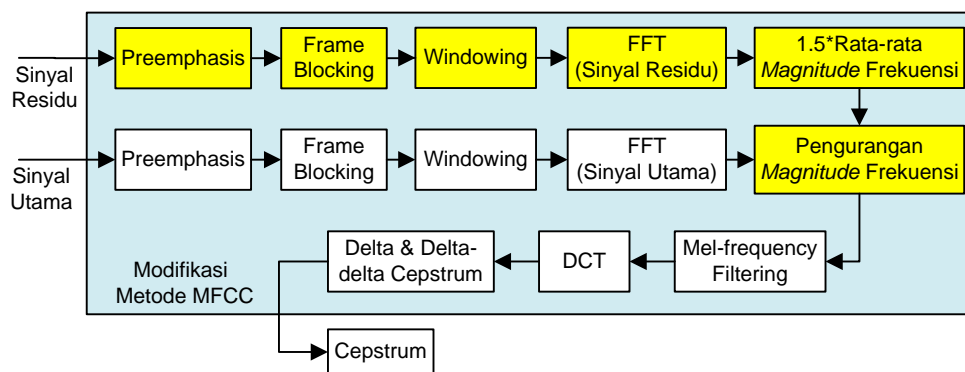
3.2.2 Modifikasi Metode MFCC

Sinyal suara setelah melewati deteksi *endpoint* akan terbagi menjadi dua bagian yaitu sinyal utama (*speech*) dan sinyal residu (*nonspeech*). Sinyal residu

kemudian diekstrak untuk mendapatkan besaran frekuensi. Frekuensi sinyal residu digunakan untuk mengurangi frekuensi sinyal utama. Blok proses pada bagian modifikasi metode MFCC ditunjukkan pada Gambar 3.2.

Tabel 3.1 Nilai Variabel pada Algoritma Deteksi *Endpoint*

Variabel	Nilai
Alpha	0.97
Jumlah sampel per <i>frame</i>	480
Jumlah sampel tumpang tindih	160
Threshold	rata-rata log energi



Gambar 3.2 Blok Proses pada Bagian Modifikasi Metode MFCC

Modifikasi metode MFCC ditunjukkan dengan blok warna kuning. Blok ini bertujuan untuk mengekstrak frekuensi sinyal residu dan mengurangkannya ke sinyal utama. Blok proses yang ditambahkan adalah *preemphasis*, *frame blocking*, *windowing*, FFT (sinyal residu), rata-rata besaran frekuensi, dan pengurangan frekuensi.

3.2.2.1 *Preemphasis*

Umumnya pada sinyal suara, energi dari frekuensi rendah lebih banyak dibandingkan dengan energi frekuensi tinggi. Perbedaan terjadi karena frekuensi rendah disampling dengan frekuensi sampling yang tinggi. *Preemphasis* digunakan untuk meningkatkan energi frekuensi tinggi dari sinyal suara. Cara kerja *preemphasis* hampir sama dengan *high-pass filter* yaitu untuk melewati

komponen frekuensi tinggi dari sinyal suara. Nilai parameter α yang digunakan pada penelitian ini adalah 0.97 (Han dkk, 2006).

3.2.2.2 Frame Blocking

Tujuan *frame blocking* adalah membagi sinyal suara menjadi *frame* kecil yang cukup sampel untuk mendapatkan informasi yang cukup (Gupta, 2013). Jika ukuran *frame* terlalu kecil, informasi yang didapat tidak cukup untuk dipercaya. Sebaliknya jika terlalu besar, informasi didalam *frame* sering berubah-ubah (Gupta, 2013). *Frame blocking* sinyal residu sama dengan *frame blocking* sinyal utama. Nilai parameter yang digunakan yaitu panjang *frame*, $N \approx 30$ milidetik, dan *frame shift*, $M \approx 10$ milidetik. Frekuensi sampel yang digunakan adalah 16000 Hz sehingga nilai $N = 480$ sampel dan $M = 160$ sampel.

3.2.2.3 Windowing

Windowing digunakan untuk meminimalkan ketidaksinambungan pada awal dan akhir *frame* (Leon, 2009). Dengan menggunakan window, sinyal suara pada awal dan akhir *frame* menjadi runcing (Gupta, 2013). Pada penelitian ini fungsi *window* yang digunakan adalah *Hamming window* yang dinyatakan dengan persamaan 2.4. Operasi ini dilakukan per *frame* sebanyak *frame* yang dihasilkan oleh proses *frame blocking*.

3.2.2.4 Fast Fourier Transform (FFT)

Fast Fourier Transform (FFT) digunakan untuk merubah sinyal suara dari domain waktu ke domain frekuensi (Gupta, 2013). Hasil dari FFT adalah *magnitude* frekuensi dalam bentuk bilangan kompleks. Kemudian bilangan kompleks yang terdiri dari riil dan imajiner dirubah ke bilangan riil dengan menghitung nilai mutlak *magnitude* bilangan kompleks. Merubah sinyal suara dari domain waktu ke domain frekuensi digunakan persamaan 2.5. Pada penelitian ini panjang FFT yang dipakai adalah 1024. Karena 1024 adalah nilai dua kali 2^n terkecil yang lebih dari nilai N (panjang *frame*).

3.2.2.5 Rata-rata *Magnitude* Frekuensi

Rata-rata *magnitude* frekuensi adalah rata-rata *magnitude* frekuensi sinyal residu pada semua *frame*. Dari proses *frame blocking* dihasilkan beberapa *frame* yang mana setiap *frame* berisi nilai *magnitude* setiap frekuensi. Nilai rata-rata yang dimaksud adalah rata-rata untuk setiap frekuensi dari semua *frame* yang sudah dihasilkan. Misalkan dari proses *frame blocking* dihasilkan 3 *frame* yaitu $F1$, $F2$, dan $F3$. Kemudian dilakukan operasi FFT untuk merubah dari domain waktu ke domain frekuensi. Didalam $F1$ dihasilkan vektor frekuensi dengan panjang D dimensi. Rata-rata *magnitude* frekuensi didapatkan dengan menjumlahkan vektor $F1$, $F2$, dan $F3$ dibagi jumlah *frame* dalam contoh ini adalah 3. Rata-rata digunakan untuk mendapatkan estimasi *magnitude* frekuensi derau.

Magnitude sinyal yang dirata-rata adalah sinyal residu hasil dari proses deteksi *endpoint*. Karena sinyal residu letaknya diawal atau diakhir dari sinyal suara, maka *magnitude* yang didapatkan tidak penuh seperti *magnitude* sinyal suara yang berada ditengah. Ketika diawal, sinyal suara cenderung naik sedangkan diakhir sinyal suara cenderung turun sehingga mengurangi *magnitude* yang dihasilkan. Oleh karena itu, pada proses rata-rata *magnitude* frekuensi terdapat nilai pengalih. Nilai pengalih yang digunakan adalah 0.5, 1, 1.5, 2, dan 2.5. Nilai pengalih didapatkan dari uji coba pada sampel data dan didapatkan nilai terbaik adalah 1.5.

3.2.2.6 Pengurangan *Magnitude* Frekuensi

Setelah didapatkan rata-rata *magnitude* frekuensi derau, selanjutnya *magnitude* frekuensi dari sinyal utama dikurangi *magnitude* frekuensi dari sinyal residu. Hasil dari proses ini adalah *magnitude* frekuensi yang bebas derau. Proses ini dilakukan untuk semua *frame* pada sinyal utama.

3.2.2.7 *Mel Frequency Filtering*

Mel frequency filtering digunakan untuk memfilter frekuensi yang dihasilkan oleh proses FFT. Pada penelitian ini nilai frekuensi batas bawah dan batas atas yang digunakan adalah 300 Hz dan 4000 Hz. Digunakan frekuensi

antara 300 Hz sampai 4000 Hz karena sebagian besar energi suara manusia terkonsentrasi antara 300 Hz sampai 3400 Hz (Nortel Networks, 2002). Sedangkan frekuensi dibawah 300 Hz dan diatas 4000 Hz tidak dipakai.

3.2.2.8 Discrete Cosine Transform (DCT)

Discrete Cosine Transform (DCT) adalah proses yang digunakan untuk mengembalikan dari domain frekuensi ke domain waktu (Leon, 2009). DCT selain untuk mengembalikan ke domain waktu juga untuk kompresi spektrum (Gupta, 2013). Pada proses DCT terdapat variabel N yaitu jumlah *cepstrum* DCT. Jumlah *cepstrum* DCT atau N pada penelitian ini digunakan nilai 13. Untuk *cepstrum* ke-0 tidak dipakai karena mengandung sedikit informasi (Leon, 2009).

3.2.2.9 Delta & Delta-delta Cepstrum

Turunan pertama (*delta*) dan turunan kedua (*delta-delta*) menggambarkan karakteristik dinamik dari sinyal suara. Nilai *delta cepstrum* didapatkan dari turunan pertama koefisien *cepstrum* sedangkan *delta-delta* didapatkan dari turunan pertama dari *delta cepstrum*. Pada proses delta & delta-delta cepstrum terdapat variabel jumlah *delta* atau T . Nilai variabel T yang digunakan pada penelitian ini adalah 2.

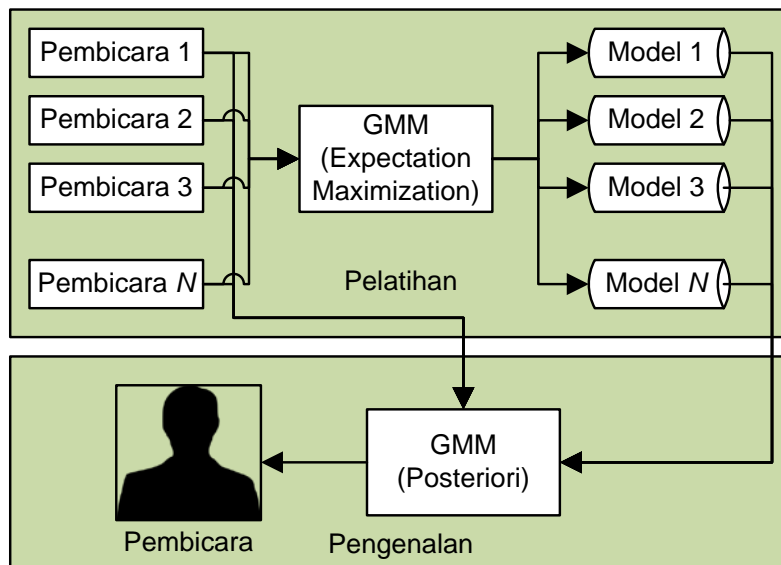
3.2.3 Pelatihan dan Pengenalan

Pelatihan dan pengenalan pada sistem ini digunakan *gaussian mixture model* (GMM). *Gaussian Mixture Model* adalah model statistik yang umum digunakan pada pengenalan pembicara dan pengenalan suara. Pada GMM terdapat variabel jumlah komponen atau M . Nilai variabel M yang digunakan pada penelitian ini adalah 32 komponen. Model pelatihan dan pengenalan menggunakan GMM ditunjukkan pada Gambar 3.3.

3.3 Rancangan Uji Coba

Skenario uji coba bertujuan untuk mendapatkan akurasi sistem idenfikasi pembicara menggunakan metode yang diusulkan, metode MFCC, metode MFCC+SS, dan metode MFCC+wiener. Hasil akurasi dari uji coba kemudian

dianalisis berdasarkan akurasi metode ekstraksi fitur terhadap pembicara, akurasi metode ekstraksi fitur terhadap tipe derau, dan akurasi metode ekstraksi fitur terhadap tingkat SNR. *Dataset* yang digunakan ada dua yaitu *dataset* sinyal suara dan *dataset* sinyal derau. Setiap pengujian dilakukan dengan tingkat SNR dan tipe derau yang berbeda. Tipe derau yang digunakan ada 5 sedangkan tingkat SNR yang digunakan ada 7.



Gambar 3.3 Pelatihan dan Pengenalan Menggunakan GMM

3.3.1 Data Uji Coba

Dataset yang digunakan dalam penelitian ini ada dua yaitu *dataset* sinyal suara dan *dataset* sinyal derau. *Dataset* sinyal suara yang digunakan diambil dari database CHAINS corpus (Cummins dkk, 2006). *Dataset* derau diambil dari database *Signal Processing Information Base* (SPIB).

Dataset sinyal suara yang dipakai adalah *dataset solo condition* yang terdiri dari 36 pembicara, 16 perempuan dan 20 laki-laki. Frekuensi sampel dari *dataset* sinyal suara dirubah dari 44100 Hz ke 16000 Hz. Kemudian dilakukan normalisasi pada *dataset* pembicara dan derau sehingga amplitudo dari sinyal bernilai antara -1 sampai 1 .

Setiap pembicara terdiri dari beberapa teks dan setiap teks terdiri dari satu file. Teks yang digunakan untuk uji coba adalah “*Play in the street up*

ahead". Teks kemudian duplikasi sebanyak 10 dengan tempo -50% , -40% , -30% , -20% , -10% , $+10\%$, $+20\%$, $+30\%$, $+40\%$, dan $+50\%$. Total semua ada 11 file untuk setiap pembicara dengan teks yang sama. Aplikasi yang digunakan untuk mengganti tempo sinyal suara adalah aplikasi Audacity.

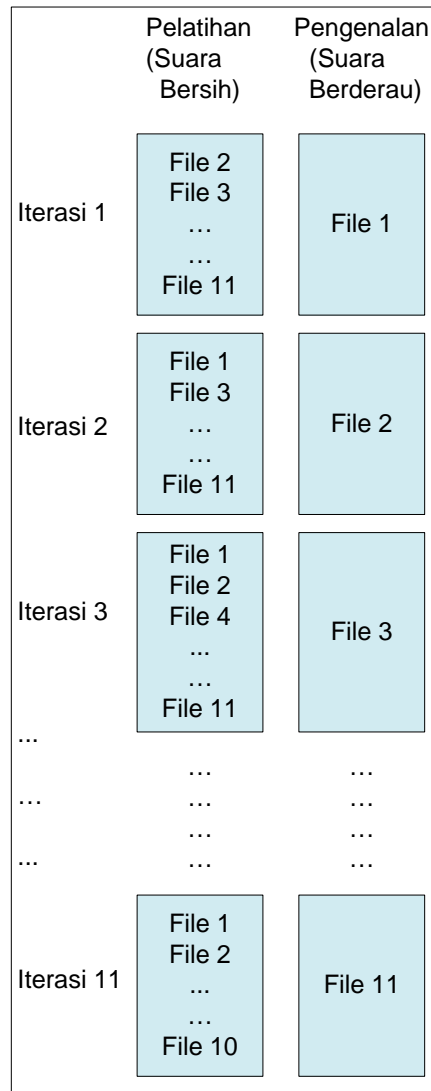
Tipe derau yang digunakan adalah *white noise*, *pink noise*, *cockpit noise* 3 (F-16), *HF channel noise*, *vehicle interior noise* (volvo 340). Derau akan ditambahkan ke dalam *dataset* sinyal suara dengan tingkat SNR bersih, 25, 20, 15, 10, 5, dan 0 dB. Setiap *dataset* ditambahi satu tipe derau dengan satu tingkat SNR. Karena ada 5 tipe derau dan 7 tingkat SNR sehingga *dataset* akan duplikasi sebanyak 35 *dataset*. Kemudian setiap dataset dilakukan normalisasi sehingga nilai amplitudo dari sinyal bernilai antara -1 sampai 1 .

3.3.2 Pengujian dan Evaluasi

Pengujian dilakukan pada tipe derau dan tingkat SNR yang berbeda. Tipe derau yang digunakan ada 5 yaitu *f16*, *hfchannel*, *pink*, *volvo*, dan *white*. Tingkat SNR yang digunakan adalah bersih, 25, 20, 15, 10, 5, dan 0 dB. Karena ada 5 tipe derau dan 7 tingkat SNR sehingga *dataset* akan duplikasi sebanyak 35 *dataset*.

Satu *dataset* terdiri dari 36 pembicara dan setiap pembicara terdiri dari 11 file. Pelatihan menggunakan 10 file pada setiap pembicara dari *dataset* sinyal yang bersih. Pengenalan menggunakan 1 file pada setiap pembicara dari *dataset* sinyal yang berderau. Pelatihan dan pengenalan untuk setiap file diulangi sebanyak 5 kali untuk mendapatkan akurasi maksimum. Uji coba ini diulangi sebanyak jumlah file pada pembicara yaitu 11 file untuk mendapatkan nilai rata-rata akurasi. Jika file nomor 1 dari setiap pembicara pada *dataset* sinyal yang berderau digunakan untuk pengenalan maka file nomor 2 sampai 11 dari setiap pembicara pada *dataset* sinyal yang bersih digunakan untuk pelatihan. Kemudian diulangi sebanyak 5 kali dan dicari nilai akurasi maksimum. Uji coba diulangi lagi untuk file nomor 2 dari setiap pembicara pada *dataset* sinyal yang berderau sebagai data pengenalan dan file nomor 1, 3, sampai 11 dari setiap pembicara pada *dataset* yang bersih sebagai data pelatihan. Begitu seterusnya sampai file nomor 11 sebagai data pengenalan dan file nomor 1 sampai 10 sebagai data

pelatihan. Dari hasil uji coba tersebut diambil nilai rata-rata akurasi. Ilustrasi dari uji coba ketika pelatihan dan pengenalan ditunjukkan pada Gambar 3.4.



Gambar 3.4 Skenario Uji Coba ketika Pelatihan dan Pengenalan

Evaluasi yang digunakan pada uji coba skenario pertama adalah dengan membandingkan kinerja metode yang diusulkan dengan MFCC standar, MFCC+SS, dan MFCC+Wiener Filter. Kinerja yang dibandingkan berdasarkan analisis akurasi setiap pembicara, akurasi setiap tipe derau, dan akurasi setiap tingkat SNR. *Dataset* uji coba yang digunakan harus sama antara metode yang

diusulkan dengan metode MFCC, MFCC+SS, dan MFCC+Wiener Filter. Akurasi identifikasi pembicara dihitung menggunakan persamaan 3.1.

$$Akurasi (\%) = \frac{Jumlah\ identifikasi\ benar}{Jumlah\ data\ uji\ coba} \times 100 , \quad (3.1)$$

BAB 4

HASIL DAN PEMBAHASAN

Pada bab ini akan dijelaskan tahapan pengolahan dataset suara dan derau, implementasi sistem identifikasi pembicara, pengujian, dan analisis hasil uji coba sistem identifikasi pembicara.

4.1 Pengolahan *Dataset*

Dataset yang digunakan pada penelitian ini ada dua yaitu *dataset* pembicara dan *dataset* derau. Kedua *dataset* tersebut didapatkan dari database yang sudah ada di internet. *Dataset* pembicara diperoleh dari database CHAINS corpus dengan alamat web <http://chains.ucd.ie/corpus.php>. Sedangkan *dataset* derau diperoleh dari database *Signal Processing Information Base* (SPIB) dengan alamat web <http://spib.linse.ufsc.br>.

CHAINS corpus adalah database pembicara yang terdiri dari 36 pembicara. Sebanyak 28 pembicara (14 laki-laki, 14 perempuan) dari Ireland dan 8 pembicara (4 laki-laki, 4 perempuan) dari UK dan USA. Informasi jenis kelamin pembicara dapat dilihat di Lampiran 1. Pada CHAINS corpus terdapat beberapa kondisi berbicara yaitu berbicara sendiri (*solo speech*), *retelling*, berbicara bersamaan (*synchronous speech*), berbicara cepat (*fast speech*), dan berbicara berbisik (*whispered speech*). Kondisi berbicara yang dipakai pada penelitian ini adalah yang kondisi berbicara sendiri (*solo speech*). Teks yang digunakan pada CHAINS corpus adalah teks dongeng pendek (*Cinderella*, *Rainbow text*, *North Wind and the Sun*, *Members of the Body*) dan 33 kalimat tunggal. Setiap teks terdiri dari satu file. Teks yang digunakan pada penelitian ini adalah kalimat tunggal nomor 3 dengan teks “*play in the street up ahead*”. Frekuensi sampling dari database CHAINS corpus adalah 44100 Hz.

Pada *dataset* derau terdiri dari beberapa tipe derau yaitu *white noise*, *pink noise*, *speech babble*, *factory floor noise 1*, *factory floor noise 2*, *cockpit noise 1* (Buccaneer Jet pada kecepatan 190 knots), *cockpit noise 2* (Buccaneer Jet pada kecepatan 450 knots), *cockpit noise 3* (F16), *engine room noise*, *operation room*

background noise, *military vehicle noise* (Leopard), *military vehicle noise* (M109), *machine gun noise*, *vehicle interior noise* (Volvo 340), dan *HF channel noise*. Tipe derau yang digunakan pada penelitian ini adalah derau tipe *white noise*, *pink noise*, *cockpit noise 3* (F16), *vehicle interior noise* (Volvo 340), dan *HF channel noise*. Frekuensi sampling pada *dataset* derau adalah 16000 Hz.

Lama waktu yang dihasilkan manusia ketika berbicara satu kalimat untuk kali kedua atau lebih tidak pernah sama meskipun untuk kalimat yang sama. Untuk menirukan kondisi seperti itu perlu dilakukan duplikasi dengan modifikasi. Duplikasi dengan modifikasi pada penelitian ini adalah memperbanyak file teks dengan tempo yang berbeda-beda. Dari satu file teks kemudian diduplikasi dengan modifikasi sebanyak 10 file dengan tempo -50%, -40%, -30%, -20%, -10%, +10%, +20%, +30%, +40%, dan +50%. Nilai tempo positif menunjukkan sinyal suara menjadi lebih pendek dan nilai negatif menunjukkan sinyal suara menjadi lebih panjang. Perubahan waktu pada sinyal suara setelah dilakukan pengubahan tempo ditunjukkan pada Lampiran 1. Aplikasi yang digunakan untuk duplikasi dengan modifikasi file suara adalah Audacity. Dalam satu pembicara jumlah file menjadi 11. Karena jumlah pembicara ada 36 pembicara sehingga total dalam satu dataset sebanyak 396 file.

Dataset pembicara dan dataset derau diolah untuk pembuatan dataset sinyal suara bersih dan dataset sinyal suara berderau. Pembuatan dataset sinyal suara berderau dengan cara menambahkan sinyal derau ke sinyal suara yang bersih. Dalam satu dataset sinyal suara yang bersih hanya ditambahi satu tipe derau dengan satu tingkat SNR. Untuk membuat dataset sinyal suara berderau dengan tipe derau dan tingkat SNR yang lain perlu dilakukan duplikasi dataset sinyal suara yang bersih kemudian ditambahi sinyal derau dengan tipe derau dan tingkat SNR yang berbeda. Tingkat SNR yang digunakan adalah bersih, 25, 20, 15, 10, 5, dan 0 dB. Karena frekuensi *sampling dataset* pembicara 44100 Hz dan *dataset* derau 16000 Hz maka perlu dilakukan penyamaan frekuensi *sampling* menjadi 16000 Hz. Kemudian amplitudo dari setiap *dataset* yang terbentuk dinormalisasi sehingga amplitudo sinyal bernilai antara -1 sampai 1.

4.2 Implementasi Sistem Identifikasi Pembicara

Pada sub bab ini akan dijelaskan tentang spesifikasi sistem yang akan digunakan sebagai ruang lingkup dan kode sumber yang akan diimplementasikan pada sistem identifikasi pembicara.

4.2.1 Spesifikasi Sistem

Pada penelitian ini, sistem identifikasi pembicara disimulasikan menggunakan perangkat keras dan perangkat lunak. Spesifikasi perangkat keras dan perangkat lunak yang digunakan sebagai berikut.

1. Perangkat Keras

- a. Prosesor : Intel® Core™ i7-3632QM @ 2.2 GHz
- b. Memory (RAM) : 4 GB 1600 MHz DDR3
- c. Tipe sistem: 64-bit sistem operasi

2. Perangkat Lunak

- a. Sistem Operasi : Windows 8.1 Professional x64 based
- b. Perangkat Pengembang : Matlab R2012a

4.2.2 Deteksi *Endpoint*

Deteksi *endpoint* digunakan untuk memisahkan antara bagian *speech* dan *nonspeech* pada sinyal suara. Pada penelitian ini, *nonspeech* dianggap berada diawal, diakhir, atau keduanya. Deteksi *endpoint* yang digunakan pada penelitian ini adalah *short-time energy*. Pada deteksi *endpoint* terdapat parameter *alpha*, panjang *frame*, *frame shift*, dan *threshold* yang berturut-turut bernilai 0.97, 30 milidetik (480 sampel), 10 milidetik (160 sampel), dan rata-rata log energi. Terdapat dua parameter input yaitu *signal_suara* dan *Fs*. *signal_suara* adalah array atau vektor bertipe double dan *Fs* adalah skalar yang mempresentasikan nilai frekuensi sampel sinyal suara. Sedangkan parameter output adalah *x_first* dan *x_last* yang merupakan posisi perpotongan antara *speech* dan *nonspeech*. Implementasi fungsi deteksi *endpoint* di Matlab ditunjukkan pada Kode Sumber 4.1.

Kode Sumber 4.1 Deteksi *Endpoint*

```
% deteksi endpoint menggunakan short-time energy
function [x_first, x_last] =
    endpoint_detection(sinyal_suara, Fs)

alpha = 0.97;    % alpha
sinyal_suara = filter([1, -alpha], 1, sinyal_suara);

frame_len = 0.03 * Fs;    % 30 milisecond
frame_shift = 0.01 * Fs;    % 10 milisecond
frame = buffer(sinyal_suara, frame_len, frame_shift,
    'nodelay');

frame = frame.^2;
frame = sum(frame);
frame = max(frame, 1e-30);
frame = 20*log10(frame);    % dalam satuan dB

threshold = mean(frame);
t_first = find(frame > threshold, 1, 'first');
t_last = find(frame > threshold, 1, 'last');

if t_first <= 1
    x_first = 1;
else
    x_first = (t_first-1)*(frame_len-frame_shift)+
        round(frame_shift/2)+1;
end
if t_last >= length(frame)
    x_last = length(sinyal_suara);
else
    x_last = (t_last)*(frame_len-frame_shift)+
        round(frame_shift/2);
end
end
```

4.2.3 *Preemphasis*

Pada *preemphasis* energi dari frekuensi tinggi ditingkatkan agar sama dengan energi frekuensi rendah. Perbedaan ini terjadi karena frekuensi rendah disampling dengan frekuensi sampling yang cukup tinggi. Nilai parameter *alpha* pada *preemphasis* bernilai 0.97. Nilai ini yang sering digunakan pada proses *preemphasis* sinyal suara. Parameter input pada proses ini adalah sinyal suara yang belum mengalami *preemphasis* sedangkan outputnya adalah sinyal suara

yang sudah di-*preemphasis*. Implementasi *preemphasis* di Matlab ditunjukkan pada Kode Sumber 4.2.

Kode Sumber 4.2 *Preemphasis*

```
% preemphasis
alpha = 0.97;          % nilai alpha
sinyal_suara = filter([1, -alpha], 1, sinyal_suara);
```

4.2.4 *Frame Blocking*

Frame blocking bertujuan untuk membagi sampel sinyal suara menjadi kelompok-kelompok kecil yang disebut *frame*. Banyaknya sampel dalam satu *frame* atau panjang *frame* yang disimpan pada variabel `frame_len` adalah 30 milidetik atau 480 sampel. Selain itu, banyaknya sampel yang tumpang tindih atau `frame_shift` dalam pengambilan sampel dari sinyal suara adalah 10 milidetik atau 160 sampel. Output dari proses ini adalah *frame* dengan panjang yang sama. Implementasi *frame blocking* di Matlab ditunjukkan pada Kode Sumber 4.3.

Kode Sumber 4.3 *Frame Blocking*

```
% frame blocking
frame = buffer(sinyal_suara, frame_len, frame_shift,
               'nodelay');
frame(:, ~any(frame,1)) = []; % menghilangkan nilai 0
```

4.2.5 *Windowing*

Setelah sinyal suara dipotong menjadi *frame*, pada *frame* terjadi ketidaksinambungan pada awal dan akhir sehingga perlu dilakukan *windowing*. Tujuan dari *windowing* adalah untuk meminimalkan ketidaksinambungan tersebut sehingga sinyal suara menjadi runcing pada awal dan akhir *frame*. Implementasi dari *windowing* di Matlab ditunjukkan pada Kode Sumber 4.4.

Kode Sumber 4.4 *Windowing*

```
% windowing
[~,c] = size(frame);
h = hamming(frame_len);
h = repmat(h, 1, c);
frame = h .* frame;
```

4.2.6 Fast Fourier Transform (FFT)

FFT digunakan untuk mengubah sinyal suara dari domain waktu ke domain frekuensi. Proses FFT dilakukan pada semua *frame* yang sudah dihasilkan oleh proses sebelumnya. Setiap satu *frame* menghasilkan satu vektor frekuensi. Pada FFT terdapat variabel panjang FFT yang disimpan pada variabel *NFFT* dengan nilai 1024. Output dari proses FFT adalah bilangan kompleks yang menyatakan frekuensi. Implementasi FFT di Matlab ditunjukkan pada Kode Sumber 4.5.

Kode Sumber 4.5 *Fast Fourier Transform (FFT)*

```
% FFT
NFFT = 1024;      % panjang FFT
X = fft(frame, NFFT);
```

4.2.7 Rata-rata *Magnitude* Frekuensi

Rata-rata *magnitude* frekuensi bertujuan untuk mendapatkan *magnitude* yang tepat, tidak terlalu besar dan tidak terlalu kecil. Input dari proses ini adalah *magnitude* sinyal residu sedangkan outputnya adalah hasil rata-rata *magnitude* sinyal residu. Pada proses ini terdapat nilai pengalih rata-rata *magnitude* yang bernilai 1.5. Nilai pengalih 1.5 berarti rata-rata *magnitude* sinyal residu akan diperbesar 1.5 kali dari nilai aslinya. Implementasi rata-rata *magnitude* di Matlab ditunjukkan pada Kode Sumber 4.6.

Kode Sumber 4.6 Rata-rata *Magnitude*

```
% rata-rata magnitude
X_mag_residu = abs(X);
X_mag_residu = 1.5*mean(X_mag_residu, 2);
```

4.2.8 Pengurangan *Magnitude* Frekuensi

Setelah didapatkan *magnitude* sinyal utama/*speech* dan sinyal residu/*nonspeech*, *magnitude* sinyal utama dikurangi *magnitude* sinyal residu untuk menghilangkan derau pada sinyal utama. Kemudian nilai *magnitude* yang

kurang dari nol akan dibulatkan menjadi nol. Implementasi pengurangan *magnitude* di Matlab ditunjukkan pada Kode Sumber 4.7.

Kode Sumber 4.7 Pengurangan *Magnitude*

```
% pengurangan magnitude
[~,c] = size(frame);
X_mag = abs(X) - repmat(X_mag_residu, 1, c);
X_mag = max(X_mag, 0);
```

4.2.9 *Mel-Frequency Filtering*

Pada proses ini terdapat beberapa parameter input yaitu frekuensi sampel, panjang filter, frekuensi maksimum, frekuensi batas bawah dan batas atas. Nilai batas bawah dan batas atas yang digunakan pada *mel-frequency filtering* adalah 300 Hz dan 4000 Hz. Implementasi *mel-frequency filtering* di Matlab ditunjukkan pada Kode Sumber 4.8.

Kode Sumber 4.8 *Mel-Frequency Filtering*

```
% Mel-frequency Filtering
Fs = 16000;
M = 25;           % panjang filter
K = NFFT/2+1;     % frekuensi maksimum
hz2mel = @(hz)(2595*log10(1+hz/700));
mel2hz = @(mel)(700*10.^(mel/2595)-700);
H = trifbank(M, K, [300 4000], Fs, hz2mel, mel2hz);
D = H * X_mag(1:K,:);
```

4.2.10 *Discrete Cosine Transform (DCT)*

Discrete cosine transform (DCT) digunakan untuk mengembalikan dari domain frekuensi ke domain waktu. Nilai parameter N atau jumlah *cepstrum* adalah 13 tetapi *cepstrum* nomor satu tidak dipakai sehingga hanya ada 12 *cepstrum* dari nomor 2 sampai nomor 13. Implementasi DCT di Matlab ditunjukkan pada Kode Sumber 4.9.

4.2.11 *Delta & Delta-delta Cepstrum*

Delta & delta-delta cepstrum adalah proses untuk mendapatkan turunan pertama dan turunan kedua. Nilai parameter T/panjang delta/jumlah tetangga yang

dipakai adalah 2. Output dari proses ini adalah vektor gabungan dari *cepstrum*, turunan pertama, dan turunan kedua. Implementasi delta & delta-delta cepstrum di Matlab ditunjukkan pada Kode Sumber 4.10.

Kode Sumber 4.9 Discrete Cosine Transform (DCT)

```
% DCT
N = 13;      % jumlah koefisien dct
cepstrum = dct(D);
cepstrum = cepstrum(2:N, :);
```

Kode Sumber 4.10 Delta & Delta-delta Cepstrum

```
% Delta & Delta-delta Cepstrum
[r,c] = size(cepstrum);
T = 2;      % konstanta delta cepstrum
div = 2*sum([1:T].^2);
C_copy = [zeros(T,c); cepstrum; zeros(T,c)];
d1 = zeros(r,c);
for i=1:1:r
    d1(i,:) = sum(C_copy(i+T:i+2*T,:)-
        flipud(C_copy(i:i+T,:)))/div;
end
C_copy = [zeros(T,c); d1; zeros(T,c)];
d2 = zeros(r,c);
for i=1:1:r
    d2(i,:) = sum(C_copy(i+T:i+2*T,:)-
        flipud(C_copy(i:i+T,:)))/div;
end
cepstrum = [cepstrum; d1; d2];
```

4.3 Uji Coba Skenario Identifikasi Pembicara

Uji coba skenario ini bertujuan untuk mendapatkan akurasi identifikasi pembicara menggunakan metode yang diusulkan, metode MFCC, metode MFCC+SS, dan metode MFCC+wiener pada tipe derau dan tingkat SNR (SNR) yang berbeda. Keempat metode tersebut menggunakan *dataset* sinyal suara bersih dan dataset sinyal suara berderau yang sama.

Ketika pelatihan menggunakan *dataset* sinyal suara yang bersih. Sedangkan ketika pengenalan menggunakan *dataset* sinyal yang berderau. Tidak semua file yang ada pada pembicara digunakan untuk pelatihan dan pengenalan. Dari 11 file pada setiap pembicara hanya 10 yang digunakan ketika pelatihan. Sedangkan ketika pengenalan hanya menggunakan 1 file pada setiap pembicara.

Jika file nomor 1 dari setiap pembicara pada *dataset* sinyal suara yang berderau digunakan untuk pengenalan maka file nomor 2 sampai 11 dari setiap pembicara pada *dataset* sinyal suara yang bersih digunakan untuk pelatihan. Kemudian diulangi sebanyak 5 kali untuk file yang sama dan diambil nilai akurasi maksimum. Uji coba diulangi lagi untuk file nomor 2 dari setiap pembicara pada *dataset* sinyal suara yang berderau sebagai data pengenalan dan file nomor 1, 3, sampai 11 dari setiap pembicara pada *dataset* sinyal suara yang bersih sebagai data pelatihan. Begitu seterusnya sampai file nomor 11 sebagai data pengenalan dan file nomor 1 sampai 10 sebagai data pelatihan. Kemudian diambil nilai rata-rata untuk setiap pembicara.

Uji coba diulangi sebanyak 5 kali untuk file yang sama dikarenakan metode *Gaussian Mixture Model* (GMM) yang digunakan untuk tidak bisa menemukan global optimum. Yang ditemukan oleh GMM adalah lokal optimum. Hal ini terjadi karena ada proses random di GMM yaitu ketika pemilihan M vektor fitur sebagai komponen GMM. Variabel M adalah jumlah komponen GMM yang digunakan ketika uji coba.

4.3.1 Nilai Variabel/Parameter pada Uji Coba

Nilai variabel/parameter yang digunakan pada metode-metode tersebut ditunjukkan pada Tabel 4.1. Semua nilai parameter pada keempat metode yang ditunjukkan pada Tabel 4.1 bernilai sama kecuali untuk parameter “*initial silence*”. Metode MFCC tidak menggunakan parameter “*initial silence*”. Sedangkan metode yang diusulkan memakai *endpoint detection* sehingga nilai parameter “*initial silence*” menjadi dinamik. Nilai variabel/parameter yang digunakan pada algoritma *endpoint detection* ditunjukkan pada tabel 4.2.

Pemilihan nilai variabel jumlah komponen GMM didasarkan pada percobaan yang sudah dilakukan. Pada uji coba dilakukan percobaan menggunakan metode yang diusulkan, MFCC, MFCC+SS, dan MFCC+wiener dengan jumlah komponen GMM sebesar 8, 16, 32, dan 64. Dari hasil percobaan tersebut kemudian dilakukan uji t untuk mengetahui pada jumlah komponen GMM berapa metode yang diusulkan menghasilkan akurasi lebih tinggi dari metode pembanding.

Tabel 4.1 Nilai Variabel/Parameter pada Uji Coba

Parameter	Metode yang Diusulkan	MFCC	MFCC+ SS	MFCC+ wiener
Alpha preemphasis	0.97	0.97	0.97	0.97
Panjang frame	480	480	480	480
Frame shift	160	160	160	160
Panjang FFT	1024	1024	1024	1024
Jumlah Mel filter bank (M)	25	25	25	25
Frekuensi batas bawah Mel filter bank	300	300	300	300
Frekuensi batas atas Mel filter bank	4000	4000	4000	4000
Jumlah koefisien DCT (N)	13	13	13	13
<i>Initial silence (IS)</i>	deteksi <i>endpoint</i>	-	10	10
Jumlah komponen GMM	32	32	32	32

Tabel 4.2 Nilai Variabel/Parameter pada Algoritma *Endpoint Detection*

Parameter	Nilai
Alpha	0.97
Panjang frame	480
Frame tumpang tindih	160
Threshold	rata-rata log energi

Nilai uji t akurasi metode ekstraksi fitur terhadap pembicara, tipe derau, dan tingkat SNR dengan komponen GMM sebesar 8, 16, 32, dan 64 ditunjukkan pada Tabel 4.3, Tabel 4.4, dan Tabel 4.5. Pada Tabel 4.3, nilai t hitung ketika komponen GMM sebesar 8 kurang dari nilai t kritis sehingga kemungkinan yang bisa dipilih adalah komponen GMM 16, 32, dan 64. Pada Tabel 4.4, nilai t hitung ketika komponen GMM sebesar 8 dan 16 kurang dari nilai t kritis sehingga kemungkinan yang bisa dipilih adalah komponen GMM 32 dan 64. Sedangkan pada Tabel 4.5 nilai t hitung dari semua komponen GMM kurang dari nilai t kritis sehingga tidak ada pilihan lain kecuali memilih nilai t hitung yang paling besar

yaitu ketika komponen GMM sebesar 32. Dengan demikian komponen GMM yang bisa dipilih dari kemungkinan yang ada adalah komponen GMM sebesar 32.

Tabel 4.3 Nilai Uji t Akurasi Metode Ekstraksi Fitur Terhadap Pembicara (t kritis = 1.69)

Komponen GMM	Nilai t Hitung		
	Metode yang diusulkan dibandingkan MFCC	Metode yang diusulkan dibandingkan MFCC+SS	Metode yang diusulkan dibandingkan MFCC+wiener
8	7.989	0.018	4.526
16	10.856	2.496	3.994
32	12.382	5.089	3.201
64	11.095	4.080	2.150
Ket: Nilai t hitung kurang dari nilai t kritis			

4.3.2 Hasil Uji Coba Skenario Identifikasi Pembicara

Dari proses uji coba yang sudah dilakukan didapatkan hasil berupa tabel akurasi identifikasi pembicara menggunakan metode yang diusulkan, metode MFCC, metode MFCC+SS, dan metode MFCC+wiener. Setiap metode menghasilkan beberapa tabel akurasi untuk setiap pembicara, tipe derau, dan tingkat SNR yang berbeda-beda. Tabel akurasi hasil uji coba ditunjukkan pada Lampiran 2.

Tabel 4.4 Nilai Uji t Akurasi Metode Ekstraksi Fitur Terhadap Tipe Derau (t kritis = 2.13)

Komponen GMM	Nilai t Hitung		
	Metode yang diusulkan dibandingkan MFCC	Metode yang diusulkan dibandingkan MFCC+SS	Metode yang diusulkan dibandingkan MFCC+wiener
8	2.855	0.010	10.775
16	3.415	1.327	10.855
32	3.643	2.334	6.040
64	3.778	2.810	3.909
Ket: Nilai t hitung kurang dari nilai t kritis			

Tabel 4.5 Nilai Uji t Akurasi Metode Ekstraksi Fitur Terhadap Tingkat SNR (t kritis = 1.94)

Komponen GMM	Nilai t Hitung		
	Metode yang diusulkan dibandingkan MFCC	Metode yang diusulkan dibandingkan MFCC+SS	Metode yang diusulkan dibandingkan MFCC+wiener
8	1.884	0.032	6.214
16	1.916	1.357	5.357
32	2.004	1.401	6.432
64	1.997	1.365	5.955
Ket: Nilai t hitung kurang dari nilai t kritis			

Akurasi yang dihasilkan oleh setiap metode bervariasi. Akurasi tertinggi adalah 100% dan terendah 0% tergantung dari tipe derau dan tingkat SNR. Pada semua metode, sebagian besar akurasi 0% terjadi pada tingkat SNR 0 dB kecuali pada tipe derau volvo. Pada tipe derau volvo, akurasi yang dihasilkan selalu 100% kecuali pada metode MFCC+wiener.

Secara umum akurasi mengalami penurunan seiring turunnya tingkat SNR. Pada metode yang diusulkan akurasi mulai mengalami penurunan ketika tingkat SNR lebih rendah sama dengan 10 dB. Pada metode MFCC, penurunan akurasi terjadi ketika tingkat SNR lebih rendah sama dengan 15 dB. Metode MFCC+SS mengalami penurunan akurasi ketika tingkat SNR lebih rendah sama dengan 10 dB. Sedangkan metode MFCC+wiener mengalami penurunan akurasi ketika tingkat SNR lebih rendah sama dengan 25 dB.

4.3.3 Analisis Hasil Uji Coba Skenario Identifikasi Pembicara

Pada hasil uji coba terdapat 4 parameter yaitu metode ekstraksi fitur, pembicara, tipe derau, dan tingkat SNR. Dari keempat parameter tersebut, hasil uji coba pada Lampiran 2 kemudian disederhanakan atau dirangkum menjadi tiga tabel. Tabel tersebut dibuat untuk mempermudah proses analisis. Tiga tabel tersebut adalah tabel akurasi metode ekstraksi fitur terhadap pembicara, tabel akurasi metode ekstraksi fitur terhadap tipe derau, dan tabel akurasi metode

ekstraksi fitur terhadap tingkat SNR. Tiga tabel tersebut digunakan untuk analisis akurasi metode ekstraksi fitur terhadap pembicara, tipe derau, dan tingkat SNR.

Analisis akurasi metode ekstraksi fitur terhadap pembicara bertujuan untuk mengetahui apakah ada keterkaitan antara metode ekstraksi fitur dengan pembicara. Keterkaitan dihitung menggunakan standar deviasi. Standar deviasi diperoleh dari empat metode ekstraksi fitur pada pembicara yang sama. Standar deviasi menyatakan perbedaan atau selisih nilai akurasi setiap metode ekstraksi fitur dengan nilai rata-rata. Tingginya perbedaan nilai akurasi setiap metode pada pembicara yang sama menunjukkan bahwa salah satu atau lebih metode yang diuji mempunyai akurasi lebih tinggi dari metode yang lain pada pembicara tersebut. Sedangkan selisih nilai akurasi setiap metode rendah menunjukkan bahwa semua metode yang diuji mempunyai akurasi yang hampir sama pada pembicara tersebut. Tabel dan grafik akurasi metode ekstraksi fitur terhadap pembicara ditunjukkan pada Lampiran 3. Sedangkan tabel nilai standar deviasi akurasi metode ekstraksi fitur terhadap pembicara ditunjukkan pada Tabel 4.6.

Berdasarkan Tabel 4.6, pembicara nomor 33 memiliki standar deviasi sangat tinggi dengan akurasi tertinggi terletak pada metode MFCC+SS dan akurasi terendah pada metode MFCC+wiener. Sedangkan pembicara nomor 8 memiliki standar deviasi sangat rendah dengan akurasi tertinggi pada tiga metode yaitu metode yang diusulkan, metode MFCC+SS, dan metode MFCC+wiener dan akurasi terendah pada metode MFCC.

Dari hasil uji coba dapat dilihat bahwa tidak ada metode yang unggul pada semua pembicara. Bagaimanapun, metode yang diusulkan mayoritas lebih unggul pada kebanyakan pembicara. Dari 36 pembicara, metode yang diusulkan unggul pada 23 pembicara atau sebesar 63.89%. Jika dilihat dari jenis kelamin, metode yang diusulkan unggul sebesar 10 dari 15 wanita atau 66.67% jenis kelamin wanita dan sebesar 13 dari 21 laki-laki atau 61.9% jenis kelamin laki-laki.

Tabel 4.6 Nilai Standar Deviasi Akurasi Metode Ekstraksi Fitur Terhadap Pembicara

Pembicara	Standar Deviasi	Pembicara	Standar Deviasi	Pembicara	Standar Deviasi
1	8.28	13	4.71	25	16.9
2	5.11	14	7.63	26	8.44
3	10.75	15	5.08	27	8.77
4	8.98	16	12.18	28	10.4
5	4.01	17	10.12	29	7.85
6	13.52	18	7.23	30	7.32
7	12.25	19	1.36	31	10.04
8	0.13	20	8.89	32	6.15
9	10.48	21	12.78	33	20.51
10	0.26	22	7.83	34	9.08
11	3.43	23	12.67	35	7.94
12	7.27	24	2.90	36	7.06

Analisis berikutnya adalah analisis akurasi metode ekstraksi fitur terhadap tipe derau. Tujuan dari analisis ini adalah untuk mengetahui apakah ada keterkaitan antara metode ekstraksi fitur dengan tipe derau. Sama dengan analisis metode ekstraksi fitur terhadap pembicara, keterkaitan metode ekstraksi fitur dengan tipe derau juga dihitung menggunakan nilai standar deviasi. Nilai standar deviasi yang besar menunjukkan bahwa salah satu atau lebih metode yang diuji mempunyai akurasi lebih tinggi dari metode yang lain pada tipe derau tersebut. Sedangkan nilai standar deviasi yang kecil menunjukkan bahwa semua metode mempunyai akurasi yang hampir sama pada tipe derau tersebut. Tabel dan grafik akurasi metode ekstraksi fitur terhadap tipe derau ditunjukkan pada Tabel 4.7 dan Gambar 4.1. Sedangkan tabel standar deviasi dan rata-rata akurasi metode ekstraksi fitur terhadap tipe derau ditunjukkan pada Tabel 4.8.

Berdasarkan Tabel 4.7 akurasi terbesar yang dihasilkan oleh metode yang diusulkan sebesar 100% pada tipe derau *volvo*. Sedangkan akurasi terendah sebesar 87.59% pada tipe derau *hfchannel*. Metode yang diusulkan memiliki akurasi tertinggi pada semua tipe derau kecuali pada tipe derau *volvo* yang akurasinya sama dengan metode MFCC dan MFCC+SS. Dengan demikian dapat diambil kesimpulan bahwa metode yang diusulkan unggul pada semua tipe derau.

Berdasarkan Tabel 4.8 tipe derau *hfchannel* memiliki standar deviasi tertinggi dengan akurasi tertinggi terletak pada metode yang diusulkan dan akurasi terendah pada metode MFCC. Sedangkan tipe derau *volvo* memiliki standar deviasi terendah dengan akurasi tertinggi pada tiga metode yaitu metode yang diusulkan, metode MFCC, dan metode MFCC+SS dan akurasi terendah pada metode MFCC+wiener. Dilihat dari nilai rata-rata, tipe derau *volvo* memiliki rata-rata akurasi tertinggi. Dengan demikian tipe derau *volvo* adalah tipe derau yang paling tidak mempengaruhi akurasi metode ekstraksi fitur.

Tabel 4.7 Akurasi Metode Ekstraksi Fitur Terhadap Tipe Derau (dalam satuan %)

Tipe Derau	Metode Ekstraksi Fitur			
	Metode yang diusulkan	MFCC	MFCC+SS	MFCC+wiener
F16	96.65	77.16	94.48	90.22
Hfchannel	87.59	66.45	82	84.27
Pink	95.49	83.48	94.91	86.58
Volvo	100	100	100	95.13
White	91.74	70.89	86.36	83.26

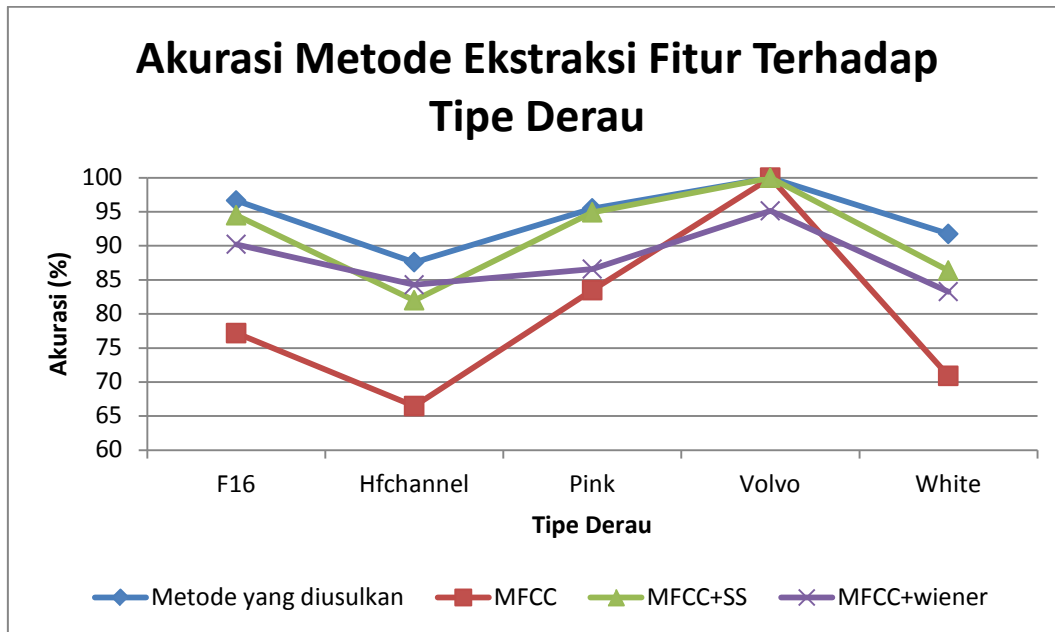
Ket: Nilai akurasi tertinggi

Tabel 4.8 Nilai Standar Deviasi dan Rata-rata Akurasi Metode Ekstraksi Fitur Terhadap Tipe Derau

Tipe Derau	Rata-rata	Standar Deviasi
F16	89.63	8.73
Hfchannel	80.08	9.37
Pink	90.12	6.01
Volvo	98.78	2.44
White	83.06	8.84

Perbedaan rentang frekuensi yang dihasilkan oleh tipe derau mempengaruhi akurasi yang dihasilkan oleh setiap metode. Setiap tipe derau menghasilkan rentang frekuensi yang berbeda-beda seperti yang ditunjukkan pada Lampiran 4. Lampiran 4 adalah gambar spektrogram dari tipe derau *f16*, *hfchannel*, *pink*, *volvo*, dan *white*. Spektrogram menggambarkan *magnitude*

frekuensi derau. Pada Gambar 4B dan 4D di Lampiran 4, tipe derau *hfchannel* menghasilkan frekuensi antara 200Hz sampai 3100Hz dengan *magnitude* sekitar -40dB sedangkan tipe derau *volvo* menghasilkan frekuensi antara 0Hz sampai 250Hz dengan *magnitude* sekitar -40dB.



Gambar 4.1 Grafik Akurasi Metode Ekstraksi Fitur Terhadap Tipe Derau

Parameter frekuensi batas bawah dan batas atas pada Mel Filter Bank yang digunakan oleh metode yang diusulkan, MFCC, MFCC+SS, dan MFCC+wiener adalah 300hz dan 4000Hz. Parameter tersebut membatasi frekuensi yang digunakan yaitu antara 300Hz sampai 4000Hz. Dengan demikian tipe derau *volvo* tidak akan mempengaruhi akurasi karena frekuensi yang dihasilkan tipe derau *volvo* berada dibawah 300Hz. Sedangkan tipe derau *hfchannel* sangat mempengaruhi akurasi karena frekuensi yang dihasilkan tipe derau *hfchannel* berada diantara 200Hz sampai 3100Hz dengan *magnitude* yang cukup tinggi yaitu -40dB.

Analisis yang terakhir adalah analisis akurasi metode ekstraksi fitur terhadap tingkat SNR. Tujuan dari analisis ini adalah untuk mengetahui apakah ada pengaruh antara tingkat SNR dengan metode ekstraksi fitur. Tabel dan grafik

akurasi metode ekstraksi fitur terhadap tingkat SNR ditunjukkan pada Tabel 4.9 dan Gambar 4.2.

Pada analisis ini berbeda dengan analisis sebelumnya karena pada analisis ini tidak akan dihitung keterkaitan antara metode ekstraksi fitur dengan tingkat SNR. Antara metode ekstraksi fitur dengan tingkat SNR tidak bisa dihitung keterkaitannya karena antar tingkat SNR saling berkaitan yaitu tingkat SNR tersebut saling berurutan dan tidak bisa diacak. Pada analisis ini hanya dapat diketahui pengaruh tingkat SNR terhadap nilai akurasi metode ekstraksi fitur. Pengaruh tingkat SNR terhadap metode ekstraksi fitur yaitu nilai akurasi yang dihasilkan oleh metode ekstraksi fitur semakin turun seiring turunnya SNR tingkat SNR seperti yang ditunjukkan pada Tabel 4.9.

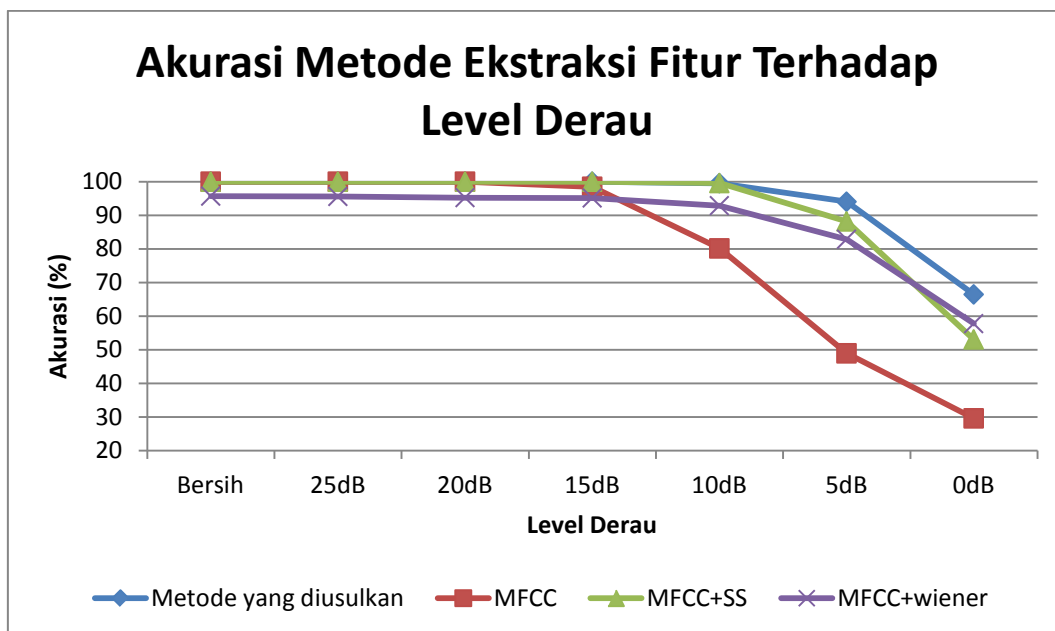
Dari Tabel 4.9 dapat dilihat bahwa nilai akurasi dari metode yang diusulkan mayoritas lebih tinggi dari metode lain kecuali pada SNR tingkat SNR 10dB. Pada SNR tingkat SNR 10dB, nilai akurasi tertinggi terletak pada metode MFCC+SS sebesar 99.6% dengan selisih 0.05% dari metode yang diusulkan. Selisih metode yang diusulkan dengan metode MFCC+SS sangat kecil sehingga akurasi kedua metode tersebut bisa dianggap sama. Dengan demikian dapat diambil kesimpulan bahwa metode yang diusulkan unggul pada semua tingkat SNR.

Tabel 4.9 Akurasi Metode Ekstraksi Fitur Terhadap Tingkat SNR (dalam satuan %)

Tingkat SNR	Metode Ekstaksi Fitur			
	Metode yang diusulkan	MFCC	MFCC+SS	MFCC+wiener
Bersih	100	100	100	95.71
25dB	100	100	100	95.61
20dB	100	100	100	95.25
15dB	100	98.48	100	95.15
10dB	99.55	80.15	99.6	92.88
5dB	94.04	48.94	88.18	82.88
0dB	66.46	29.6	53.08	57.78

Ket: Nilai akurasi tertinggi

Penurunan akurasi seiring turunnya SNR tingkat SNR disebabkan karena frekuensi pembicara tertutup oleh frekuensi derau. Frekuensi derau dengan *magnitude* lebih besar menutupi frekuensi pembicara yang mempunyai *magnitude* lebih kecil. Sebagai ilustrasi dapat dilihat gambar spektrogram dari pembicara 1 pada Lampiran 5. Pada Lampiran 5 terdapat gambar spektrogram dari pembicara 1 pada derau *hfchannel* dengan tingkat SNR bersih sampai 0dB. Ketika tingkat SNR bersih, frekuensi pembicara terlihat jelas. Kemudian frekuensi pembicara sedikit demi sedikit tersamarkan dan mirip dengan frekuensi derau ketika tingkat SNR diturunkan.



Gambar 4.2 Grafik Akurasi Metode Ekstraksi Fitur Terhadap Tingka SNR

Dari hasil analisis yang sudah dilakukan, metode yang diusulkan unggul pada semua analisis. Metode yang diusulkan unggul pada 23 pembicara dari 36 pembicara. Metode yang diusulkan juga unggul pada semua tipe derau kecuali tipe derau *volvo* yang sama dengan metode MFCC dan MFCC+SS. Pada kondisi tingkat SNR bermacam-macam, metode yang diusulkan unggul pada semua tingkat SNR kecuali pada SNR tingkat SNR 10dB yang lebih unggul metode MFCC+SS dengan selisih 0.05% dari metode yang diusulkan. Rata-rata akurasi metode yang diusulkan, MFCC, MFCC+SS, dan MFCC+wiener secara berturut-

turut adalah 94.29%, 79.6%, 91.55%, dan 87.89%. Dengan demikian dapat disimpulkan bahwa metode yang diusulkan memiliki kinerja lebih baik dari metode MFCC, MFCC+SS, dan MFCC+wiener dalam hal akurasi.

Metode yang diusulkan memiliki kinerja yang lebih baik dari metode MFCC, MFCC+SS, dan MFCC+wiener karena metode yang diusulkan menggunakan residu (sisa hasil) algoritma deteksi *endpoint* sebagai informasi frekuensi derau. Residu algoritma deteksi *endpoint* diekstrak untuk mendapatkan *magnitude* frekuensi derau. *Magnitude* frekuensi derau digunakan sebagai pengurang pada sinyal utama sehingga sinyal utama bersih dari derau.

4.3.4 Uji t Hasil Analisis Sistem Identifikasi Pembicara

Untuk membuktikan bahwa akurasi metode yang diusulkan lebih tinggi dari pada metode MFCC, MFCC+SS, dan MFCC+wiener perlu dilakukan uji statistik. Uji statistik yang dipakai pada penelitian ini adalah uji t dua kelompok berpasangan dengan alpha 5%. Hipotesis nol (H_0) yang diambil adalah akurasi metode yang diusulkan lebih kecil sama dengan metode pembanding sedangkan hipotesis satu (H_1) adalah akurasi metode yang diusulkan lebih besar dari metode pembanding. Yang dimaksud kelompok berpasangan pada penelitian ini adalah akurasi yang dihasilkan oleh satu metode ekstraksi fitur. Akurasi satu metode ekstraksi fitur atau satu kelompok kemudian dibandingkan dengan akurasi metode ekstraksi fitur atau kelompok yang lain. Perhitungan uji t dinyatakan dengan persamaan 4.1.

$$t = \frac{\bar{X}_D - \mu_0}{S_D / \sqrt{n}} \quad (4.1)$$

dimana \bar{X}_D adalah rata-rata nilai selisih, S_D adalah standar deviasi dari nilai selisih, dan n adalah jumlah sampel. Sedangkan μ_0 adalah rata-rata standar yang mana pada penelitian ini bernilai 0.

Dari proses analisi dihasilkan 3 tabel akurasi yaitu tabel akurasi metode ekstraksi fitur terhadap pembicara pada Lampiran 3, tabel akurasi metode ekstraksi fitur terhadap tipe derau pada Tabel 4.7, dan tabel akurasi metode ekstraksi fitur terhadap tingkat SNR pada Tabel 4.9. Setiap tabel tersebut akan

dilakukan uji t antara metode yang diusulkan dengan metode MFCC, MFCC+SS, dan MFCC+wiener.

Uji t pertama dilakukan pada tabel akurasi metode ekstraksi fitur terhadap pembicara yang bertujuan untuk membuktikan bahwa metode yang diusulkan memiliki akurasi yang lebih tinggi dibandingkan metode MFCC, MFCC+SS, dan MFCC+wiener pada pembicara yang berbeda-beda. Ada 3 hasil uji t yang bisa didapatkan yaitu hasil uji t metode yang diusulkan dengan MFCC, hasil uji t metode yang diusulkan dengan MFCC+SS, dan hasil uji t metode yang diusulkan dengan MFCC+wiener. Hasil t hitung antara metode yang diusulkan dengan metode MFCC, MFCC+SS, dan MFCC+wiener pada pembicara yang berbeda-beda secara berturut-turut adalah 12.38, 5.09, dan 3.2. Sedangkan nilai t kritis adalah 1.69 sehingga dapat diambil kesimpulan bahwa metode yang diusulkan memiliki rata-rata akurasi yang lebih tinggi dari metode MFCC, MFCC+SS, dan MFCC+wiener pada pembicara yang berbeda-beda.

Uji t berikutnya dilakukan pada tabel akurasi metode ekstraksi fitur terhadap tipe derau yang bertujuan untuk membuktikan bahwa metode yang diusulkan memiliki akurasi yang lebih tinggi dibandingkan metode MFCC, MFCC+SS, dan MFCC+wiener pada tipe derau yang berbeda-beda. Hasil t hitung antara metode yang diusulkan dengan metode MFCC, MFCC+SS, dan MFCC+wiener pada tipe derau yang berbeda-beda secara berturut-turut adalah 3.64, 2.33, dan 6.04. Sedangkan nilai t kritis adalah 2.13 sehingga dapat diambil kesimpulan bahwa metode yang diusulkan memiliki rata-rata akurasi yang lebih tinggi dari metode MFCC, MFCC+SS, dan MFCC+wiener pada tipe derau yang berbeda-beda.

Uji t terakhir dilakukan pada tabel akurasi metode ekstraksi fitur terhadap tingkat SNR yang bertujuan untuk membuktikan bahwa metode yang diusulkan memiliki akurasi yang lebih tinggi dibandingkan metode MFCC, MFCC+SS, dan MFCC+wiener pada tingkat SNR yang berbeda-beda. Hasil t hitung antara metode yang diusulkan dengan metode MFCC, MFCC+SS, dan MFCC+wiener pada tingkat SNR yang berbeda-beda secara berturut-turut adalah 2.0, 1.4, dan 6.43. Sedangkan nilai t kritis adalah 1.94 sehingga dapat diambil kesimpulan bahwa metode yang diusulkan memiliki rata-rata akurasi yang lebih tinggi dari

metode MFCC dan MFCC+wiener tetapi lebih rendah sama dengan metode MFCC+SS pada tingkat SNR yang berbeda-beda.

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan yang dapat diambil dari penelitian ini adalah sebagai berikut.

1. Sinyal residu dari algoritma deteksi *endpoint* dapat digunakan sebagai informasi frekuensi derau. Besaran (*magnitude*) frekuensi derau diperoleh dari ekstraksi sinyal residu menggunakan *fast fourier transform* (FFT).
2. Derau pada sinyal utama dapat dihilangkan menggunakan besaran frekuensi residu. Besaran frekuensi sinyal utama dikurangi secara langsung dengan besaran frekuensi residu.
3. Modifikasi metode MFCC menggunakan sinyal residu dari algoritma deteksi *endpoint* dapat meningkatkan kinerja sistem dalam hal akurasi. Metode yang diusulkan, MFCC, MFCC+SS, dan MFCC+wiener berturut-turut memiliki rata-rata akurasi 94.29%, 79.6%, 91.55%, dan 87.89%. Metode yang diusulkan menunjukkan peningkatan akurasi sebesar 14.69% dari metode MFCC, 2.74% dari MFCC+SS, dan 6.4% dari MFCC+wiener.
4. Metode yang diusulkan unggul pada mayoritas pembicara sebesar 63.89%. Jika dilihat dari jenis kelamin, metode yang diusulkan unggul sebesar 66.67% untuk jenis kelamin wanita dan sebesar 61.9% untuk jenis kelamin laki-laki.
5. Metode yang diusulkan unggul pada semua tipe derau. Tipe derau yang paling mempengaruhi akurasi adalah tipe derau *hfchannel* dengan rata-rata akurasi 80.08 dan standar deviasi 9.37. Sedangkan tipe derau yang paling tidak mempengaruhi akurasi adalah tipe derau *volvo* dengan rata-rata akurasi 98.78 dan standar deviasi 2.44.
6. Akurasi yang dihasilkan metode ekstraksi fitur turun seiring turunnya SNR tingkat SNR. Metode yang diusulkan unggul hampir pada semua tingkat SNR kecuali pada tingkat SNR 10dB. Pada tingkat SNR 10dB, metode MFCC+SS memiliki akurasi tertinggi sebesar 99.6% dengan selisih 0.05% dari metode yang diusulkan.

7. Berdasarkan hasil uji t dapat disimpulkan bahwa metode yang diusulkan memiliki akurasi lebih tinggi daripada metode MFCC, MFCC+SS, dan MFCC+wiener pada pembicara yang berbeda-beda dan tipe derau yang berbeda-beda. Selain itu metode yang diusulkan juga memiliki akurasi lebih tinggi daripada metode MFCC dan MFCC+wiener tetapi lebih rendah sama dengan metode MFCC+SS pada tingkat SNR yang berbeda-beda.

5.2 Saran

Adapun saran yang bisa diberikan berdasarkan hasil yang didapat dari penelitian ini adalah sebagai berikut.

1. Pada penelitian ini, teks yang digunakan ketika pelatihan dan pengenalan adalah teks yang sama. Untuk penelitian berikutnya perlu dilakukan penelitian menggunakan teks yang berbeda ketika pelatihan dan pengenalan.
2. Derau yang digunakan pada penelitian ini adalah derau yang bersifat *stationer*. Untuk penelitian berikutnya perlu dilakukan penelitian menggunakan derau *nonstationer*.

DAFTAR PUSTAKA

- Chauhan, P.M. dan Desai, N.P. (2014), “Mel Frequency Cepstral Coefficients (MFCC) Based Speaker Identification in Noisy Environment Using Wiener Filter”, *2014 International Conference on Green Computing Communication and Electrical Engineering*, hal. 1-5.
- Cummins, F., Grimaldi, M., Leonard, T., dan Simko, J. (2006), “The Chains Corpus: Characterizing Individual Speaker”, *International Conference Speech and Computer (SPECOM)*, hal. 431-435.
- Gupta, S., Jaafar, J., Ahmad, W.F., dan Bansal, A. (2013), “Feature Extraction Using MFCC”, *Signal & Image Processing: An International Journal (SIPIJ)*, Vol. 4, No. 4, hal. 101-108.
- Han, W., Chan, C.F., Choy, C.S., dan Pun, K.P. (2006), “An Efficient MFCC Extraction Method in Speech Recognition”, *2006 IEEE International Symposium on Circuits and Systems*, doi: 10.1109/ISCAS.2006.1692543.
- Junqin, W. dan Junjun, Y. (2011), “An Improved Arithmetic of MFCC in Speech Recognition System”, *2011 International Conference on Electronics, Communication and Control (ICECC)*, hal. 719-722.
- Kinnunen, T., Saeidi, R., Sedlak, F., Lee, K.A., Sandberg, J., Sandsten, M.H., dan Li, H. (2012), “Low-Variance Multitaper MFCC Feature: A Case Study in Robust Speaker Verification”, *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 20, No. 7, hal. 1990-2001.
- Lee, S.M., Fang, S.H., Hung, J.W., dan Lee, L.S. (2002), “Improved MFCC Feature Extraction by PCA-Optimized Filter Bank for Speech Recognition”, *IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU'01)*, hal. 49-52.
- Leon, C.G.K (2009), “Robust Computer Voice Recognition Using Improved MFCC Algorithm”, *2009 International Conference on New Trends in Information and Service Science*, hal. 835-840.

- Lim, J.S. dan Oppenheim, A.V. (1979), "Enhancement and Bandwidth Compression of Noisy Speech", *Proceedings of The IEEE*, Vol. 67, No. 12, hal. 1586-1604.
- Matsumoto, K., Hayasaka, N., dan Iiguni, Y. (2014), "Noise Robust Speaker Identification by Dividing MFCC", 6th International Symposium on Communication, Control and Signal Processing (ISCCSP), hal. 652-655.
- Nortel Networks. (2002), *Voice Fundamentals Reference Guide*, Nortel Network Corporation, Canada.
- Reynolds, D.A. dan Rose, R.C. (1995), "Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models", *IEEE Transactions on Speech and Audio Processing*, Vol. 3, No. 1, hal. 72-83.
- Skowronski, M. D. dan Harris, J.G. (2002), "Increased MFCC Filter Bandwidth for Noise-Robust Phoneme Recognition", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vol. 1, hal. 801-804.
- Vergin, R. dan O'Shaughnessy, D. (1995), "Pre-Emphasis and Speech Recognition", *Canadian Conference on Electrical and Computer Engineering*, Vol. 2, hal. 1062-1065.
- Wu, Bing-Fei dan Wang, Kun-Ching. (2005), "Robust Endpoint Detection Algorithm Based on the Adaptive Band-Partitioning Spectral Entropy in Adverse Environments", *IEEE Transactions on Speech and Audio Processing*, Vol. 13, No. 5, hal. 762-775.
- Yong, S. dan Leimin, C. (2011), "Performance Comparison of New Endpoint Detection Method in Noise Environments", *2011 International Conference on Electric Information and Control Engineering (ICEICE)*, hal. 1523-1527.
- Young, S. (1996), "A Review of Large-Vocabulary Continuous-Speech Recognition", *IEEE Signal Processing Magazine*, Vol. 13, No. 5.
- Yutai, W., Bo, L., Xiaoqing, J., Feng, L., dan Lihao, W. (2009), "Speaker Recognition Based on Dynamic MFCC Parameters", *International Conference on Image Analysis and Signal Processing (IASP 2009)*, hal. 406-409.

- Zhang, Y. dan Abdulla, W.H. (2007), "Robust Speaker Identification in Noisy Environment using Cross Diagonal GTF-ICA Feature", *6th International Conference on Information, Communication & Signal Processing*, hal. 1-4.
- Zunjing, Wu dan Zhigang, Cao. (2005), "Improved MFCC-Based Feature for Robust Speaker Identification", *Tsinghua Science and Technology*, Vol. 10, No. 2, hal. 158-161.

LAMPIRAN 1

Jenis Kelamin Pembicara dan Panjang Sinyal Suara Pembicara Dalam Satuan Milidetik

Pembicara	Jenis Kelamin	Tempo (%)										
		-50	-40	-30	-20	-10	0	10	20	30	40	50
1	P	2731	2311	2031	1751	1611	1437	1330	1190	1120	1050	980
2	P	2941	2451	2171	1891	1681	1537	1401	1260	1190	1120	1050
3	L	2451	2101	1821	1611	1401	1288	1190	1050	980	910	840
4	P	2941	2521	2171	1891	1681	1538	1190	1120	980	910	910
5	L	2521	2101	1821	1611	1471	1310	1190	1120	980	910	910
6	L	2801	2381	2031	1821	1611	1458	1330	1190	1120	1050	980
7	L	2661	2241	1961	1751	1541	1403	1260	1190	1050	980	910
8	L	2801	2311	2031	1751	1611	1440	1330	1190	1120	1050	980
9	P	4341	3641	3151	2731	2451	2234	2031	1891	1751	1611	1471
10	P	2604	5112	4271	3641	3221	2871	2381	2171	2031	1891	1751
11	P	3081	2591	2241	1961	1751	1597	1471	1330	1260	1120	1050
12	P	3571	3011	2591	2311	2031	1859	1681	1541	1401	1330	1260
13	P	3011	2521	2171	1961	1751	1571	1401	1330	1190	1120	1050
14	P	3641	3081	2661	2311	2101	1880	1681	1541	1471	1330	1260
15	P	3431	2871	2521	2171	1961	1788	1611	1471	1401	1260	1190
16	P	2731	2311	1961	1751	1541	1424	1260	1190	1120	1050	980
17	P	3081	2591	2241	1961	1751	1593	1471	1330	1260	1120	1050
18	P	2661	2241	1961	1681	1541	1385	1260	1190	1050	980	910
19	P	3221	2731	2381	2101	1821	1679	1541	1401	1330	1190	1120
20	P	2731	2241	1961	1751	1541	1409	1260	1190	1120	980	980
21	L	2241	1891	1611	1401	1260	1159	1051	980	910	840	770
22	L	2311	1961	1681	1470	1330	1227	1120	1050	980	910	840
23	L	3501	2941	2521	2241	2031	1816	1681	1541	1401	1330	1190
24	L	2451	2031	1751	1541	1401	1278	1190	1050	980	910	840
25	L	3081	2591	2241	1961	1751	1597	1471	1330	1260	1120	1050
26	L	2311	1961	1681	1471	1330	1208	1120	980	910	840	840
27	L	2451	2031	1751	1541	1401	1281	1190	1050	980	910	840
28	L	2941	2521	2171	1891	1681	1541	1401	1260	1190	1120	1050
29	L	2521	2101	1821	1611	1401	1300	1190	1120	980	910	910
30	L	2871	2381	2101	1821	1611	1488	1330	1260	1120	1050	980
31	L	3151	2661	2311	2031	1821	1649	1471	1401	1260	1190	1120
32	L	2521	2101	1821	1611	1401	1300	1190	1120	980	910	910
33	L	2801	2381	2031	1821	1611	1456	1330	1190	1120	1050	980
34	L	2661	2241	1961	1681	1541	1384	1260	1190	1050	980	910
35	L	2801	2381	2031	1821	1611	1465	1330	1190	1120	1050	980
36	L	3501	2941	2521	2241	2031	1812	1681	1541	1401	1330	1190

Ket: L=laki-laki, P=perempuan

[Halaman ini sengaja dikosongkan]

LAMPIRAN 2

Hasil Uji Coba Sistem Identifikasi Pembicara

2.A. Akurasi Metode yang Diusulkan pada Derau F16 (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	100
2	100	100	100	100	100	90.91	27.27
3	100	100	100	100	100	100	72.73
4	100	100	100	100	100	100	100
5	100	100	100	100	100	100	100
6	100	100	100	100	100	100	54.55
7	100	100	100	100	100	100	72.73
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100
11	100	100	100	100	100	100	45.45
12	100	100	100	100	100	100	100
13	100	100	100	100	100	100	27.27
14	100	100	100	100	100	100	81.82
15	100	100	100	100	100	100	100
16	100	100	100	100	100	100	100
17	100	100	100	100	100	100	100
18	100	100	100	100	100	90.91	72.73
19	100	100	100	100	100	100	100
20	100	100	100	100	100	100	90.91
21	100	100	100	100	100	100	36.36
22	100	100	100	100	100	100	36.36
23	100	100	100	100	100	100	100
24	100	100	100	100	100	100	54.55
25	100	100	100	100	100	100	63.64
26	100	100	100	100	100	81.82	0
27	100	100	100	100	100	100	18.18
28	100	100	100	100	100	100	90.91
29	100	100	100	100	100	100	90.91
30	100	100	100	100	100	100	100
31	100	100	100	100	100	100	100
32	100	100	100	100	100	100	100
33	100	100	100	100	100	100	100

34	100	100	100	100	100	100	100
35	100	100	100	100	100	100	54.55
36	100	100	100	100	100	100	100

2.B. Metode yang Diusulkan pada Derau Hfchannel (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	100
2	100	100	100	100	90.91	81.82	9.091
3	100	100	100	100	100	90.91	0
4	100	100	100	100	100	100	36.36
5	100	100	100	100	100	100	72.73
6	100	100	100	100	100	72.73	0
7	100	100	100	100	81.82	27.27	0
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	27.27
10	100	100	100	100	100	100	100
11	100	100	100	100	81.82	63.64	18.18
12	100	100	100	100	100	90.91	18.18
13	100	100	100	100	100	9.091	0
14	100	100	100	100	100	81.82	18.18
15	100	100	100	100	100	100	90.91
16	100	100	100	100	100	81.82	72.73
17	100	100	100	100	100	100	45.45
18	100	100	100	100	100	72.73	9.091
19	100	100	100	100	100	100	90.91
20	100	100	100	100	100	90.91	45.45
21	100	100	100	100	100	63.64	0
22	100	100	100	100	100	45.45	0
23	100	100	100	100	100	100	36.36
24	100	100	100	100	100	100	9.091
25	100	100	100	100	100	54.55	0
26	100	100	100	100	100	36.36	0
27	100	100	100	100	100	63.64	0
28	100	100	100	100	100	100	18.18
29	100	100	100	100	100	100	18.18
30	100	100	100	100	100	100	45.45
31	100	100	100	100	100	100	54.55
32	100	100	100	100	100	100	90.91
33	100	100	100	100	100	36.36	9.091
34	100	100	100	100	100	100	90.91
35	100	100	100	100	72.73	0	0

36	100	100	100	100	100	100	54.55
----	-----	-----	-----	-----	-----	-----	-------

2.C. Metode yang Diusulkan pada Derau Pink (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	100
2	100	100	100	100	100	63.64	0
3	100	100	100	100	100	100	72.73
4	100	100	100	100	100	100	100
5	100	100	100	100	100	100	100
6	100	100	100	100	100	100	90.91
7	100	100	100	100	100	90.91	0
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100
11	100	100	100	100	100	100	27.27
12	100	100	100	100	100	100	63.64
13	100	100	100	100	100	100	18.18
14	100	100	100	100	100	100	36.36
15	100	100	100	100	100	100	100
16	100	100	100	100	100	100	100
17	100	100	100	100	100	100	100
18	100	100	100	100	100	100	45.45
19	100	100	100	100	100	100	100
20	100	100	100	100	100	100	90.91
21	100	100	100	100	100	100	45.45
22	100	100	100	100	100	100	9.091
23	100	100	100	100	100	100	90.91
24	100	100	100	100	100	100	27.27
25	100	100	100	100	100	100	100
26	100	100	100	100	100	90.91	9.091
27	100	100	100	100	100	100	54.55
28	100	100	100	100	100	100	63.64
29	100	100	100	100	100	100	100
30	100	100	100	100	100	100	100
31	100	100	100	100	100	100	100
32	100	100	100	100	100	100	81.82
33	100	100	100	100	100	100	81.82
34	100	100	100	100	100	100	45.45
35	100	100	100	100	100	100	72.73
36	100	100	100	100	100	100	90.91

2.D. Metode yang Diusulkan pada Derau Volvo (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	100
2	100	100	100	100	100	100	100
3	100	100	100	100	100	100	100
4	100	100	100	100	100	100	100
5	100	100	100	100	100	100	100
6	100	100	100	100	100	100	100
7	100	100	100	100	100	100	100
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100
11	100	100	100	100	100	100	100
12	100	100	100	100	100	100	100
13	100	100	100	100	100	100	100
14	100	100	100	100	100	100	100
15	100	100	100	100	100	100	100
16	100	100	100	100	100	100	100
17	100	100	100	100	100	100	100
18	100	100	100	100	100	100	100
19	100	100	100	100	100	100	100
20	100	100	100	100	100	100	100
21	100	100	100	100	100	100	100
22	100	100	100	100	100	100	100
23	100	100	100	100	100	100	100
24	100	100	100	100	100	100	100
25	100	100	100	100	100	100	100
26	100	100	100	100	100	100	100
27	100	100	100	100	100	100	100
28	100	100	100	100	100	100	100
29	100	100	100	100	100	100	100
30	100	100	100	100	100	100	100
31	100	100	100	100	100	100	100
32	100	100	100	100	100	100	100
33	100	100	100	100	100	100	100
34	100	100	100	100	100	100	100
35	100	100	100	100	100	100	100
36	100	100	100	100	100	100	100

2.E. Metode yang Diusulkan pada Derau White (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	81.82
2	100	100	100	100	90.91	54.55	54.55
3	100	100	100	100	100	100	0
4	100	100	100	100	100	100	72.73
5	100	100	100	100	100	100	90.91
6	100	100	100	100	100	81.82	45.45
7	100	100	100	100	100	81.82	18.18
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	81.82
10	100	100	100	100	100	100	100
11	100	100	100	100	100	81.82	9.091
12	100	100	100	100	100	100	63.64
13	100	100	100	100	100	90.91	0
14	100	100	100	100	100	72.73	27.27
15	100	100	100	100	100	100	90.91
16	100	100	100	100	100	100	90.91
17	100	100	100	100	100	100	81.82
18	100	100	100	100	100	100	72.73
19	100	100	100	100	100	100	100
20	100	100	100	100	100	90.91	27.27
21	100	100	100	100	100	100	0
22	100	100	100	100	100	90.91	9.091
23	100	100	100	100	100	100	63.64
24	100	100	100	100	100	100	0
25	100	100	100	100	100	90.91	9.091
26	100	100	100	100	100	54.55	9.091
27	100	100	100	100	100	100	0
28	100	100	100	100	100	90.91	45.45
29	100	100	100	100	100	100	54.55
30	100	100	100	100	100	100	90.91
31	100	100	100	100	100	100	36.36
32	100	100	100	100	100	100	81.82
33	100	100	100	100	100	100	54.55
34	100	100	100	100	100	81.82	36.36
35	100	100	100	100	100	90.91	0
36	100	100	100	100	100	100	72.73

2.F. Metode MFCC pada Derau F16 (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	90.91	0
2	100	100	100	100	100	54.55	0
3	100	100	100	100	18.18	0	0
4	100	100	100	100	100	27.27	0
5	100	100	100	100	100	100	9.091
6	100	100	100	100	9.091	0	0
7	100	100	100	100	100	0	0
8	100	100	100	100	100	100	100
9	100	100	100	100	100	0	0
10	100	100	100	100	100	100	100
11	100	100	100	100	100	81.82	0
12	100	100	100	100	100	72.73	0
13	100	100	100	100	100	81.82	54.55
14	100	100	100	100	100	0	0
15	100	100	100	100	100	100	18.18
16	100	100	100	100	100	81.82	0
17	100	100	100	100	100	27.27	0
18	100	100	100	100	100	0	0
19	100	100	100	100	100	100	100
20	100	100	100	100	100	9.091	0
21	100	100	100	100	45.45	0	0
22	100	100	100	100	54.55	0	0
23	100	100	100	100	100	0	0
24	100	100	100	100	100	100	9.091
25	100	100	100	100	18.18	0	0
26	100	100	100	100	18.18	0	0
27	100	100	100	100	72.73	0	0
28	100	100	100	100	90.91	0	0
29	100	100	100	100	100	9.091	0
30	100	100	100	100	100	54.55	0
31	100	100	100	100	100	72.73	0
32	100	100	100	100	100	100	9.091
33	100	100	100	100	100	90.91	0
34	100	100	100	100	100	100	0
35	100	100	100	100	63.64	0	0
36	100	100	100	100	100	0	0

2.G. Metode MFCC pada Derau Hfchannel (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	0	0
2	100	100	100	100	90.91	0	0
3	100	100	100	100	0	0	0
4	100	100	100	100	0	0	0
5	100	100	100	100	100	90.91	0
6	100	100	100	54.55	0	0	0
7	100	100	100	45.45	0	0	0
8	100	100	100	100	100	100	100
9	100	100	100	100	0	0	0
10	100	100	100	100	100	100	100
11	100	100	100	100	100	9.091	0
12	100	100	100	100	36.36	0	0
13	100	100	100	100	54.55	0	0
14	100	100	100	100	0	0	0
15	100	100	100	100	100	36.36	0
16	100	100	100	100	81.82	0	0
17	100	100	100	100	18.18	0	0
18	100	100	100	100	18.18	0	0
19	100	100	100	100	100	100	90.91
20	100	100	100	100	45.45	0	0
21	100	100	100	54.55	0	0	0
22	100	100	100	81.82	0	0	0
23	100	100	100	100	100	9.091	0
24	100	100	100	100	63.64	9.091	0
25	100	100	100	45.45	0	0	0
26	100	100	100	45.45	0	0	0
27	100	100	100	100	0	0	0
28	100	100	100	100	0	0	0
29	100	100	100	100	90.91	0	0
30	100	100	100	100	100	0	0
31	100	100	100	100	0	0	0
32	100	100	100	100	100	63.64	18.18
33	100	100	100	100	18.18	0	0
34	100	100	100	100	100	45.45	0
35	100	100	100	100	0	0	0
36	100	100	100	100	100	27.27	0

2.H. Metode MFCC pada Derau Pink (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	72.73	0
2	100	100	100	100	100	100	0
3	100	100	100	100	100	0	0
4	100	100	100	100	100	90.91	0
5	100	100	100	100	100	100	90.91
6	100	100	100	100	100	9.091	0
7	100	100	100	100	100	0	0
8	100	100	100	100	100	100	100
9	100	100	100	100	100	90.91	0
10	100	100	100	100	100	100	100
11	100	100	100	100	100	54.55	0
12	100	100	100	100	100	90.91	0
13	100	100	100	100	100	100	0
14	100	100	100	100	100	81.82	0
15	100	100	100	100	100	100	100
16	100	100	100	100	100	100	0
17	100	100	100	100	100	36.36	0
18	100	100	100	100	100	63.64	0
19	100	100	100	100	100	100	100
20	100	100	100	100	100	27.27	0
21	100	100	100	100	100	0	0
22	100	100	100	100	100	0	0
23	100	100	100	100	100	100	0
24	100	100	100	100	100	100	90.91
25	100	100	100	100	100	0	0
26	100	100	100	100	100	0	0
27	100	100	100	100	100	9.091	0
28	100	100	100	100	100	18.18	0
29	100	100	100	100	100	100	0
30	100	100	100	100	100	100	0
31	100	100	100	100	100	45.45	0
32	100	100	100	100	100	100	100
33	100	100	100	100	100	100	0
34	100	100	100	100	100	100	63.64
35	100	100	100	100	100	0	0
36	100	100	100	100	100	100	0

2.I. Metode MFCC pada Derau Volvo (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	100
2	100	100	100	100	100	100	100
3	100	100	100	100	100	100	100
4	100	100	100	100	100	100	100
5	100	100	100	100	100	100	100
6	100	100	100	100	100	100	100
7	100	100	100	100	100	100	100
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100
11	100	100	100	100	100	100	100
12	100	100	100	100	100	100	100
13	100	100	100	100	100	100	100
14	100	100	100	100	100	100	100
15	100	100	100	100	100	100	100
16	100	100	100	100	100	100	100
17	100	100	100	100	100	100	100
18	100	100	100	100	100	100	100
19	100	100	100	100	100	100	100
20	100	100	100	100	100	100	100
21	100	100	100	100	100	100	100
22	100	100	100	100	100	100	100
23	100	100	100	100	100	100	100
24	100	100	100	100	100	100	100
25	100	100	100	100	100	100	100
26	100	100	100	100	100	100	100
27	100	100	100	100	100	100	100
28	100	100	100	100	100	100	100
29	100	100	100	100	100	100	100
30	100	100	100	100	100	100	100
31	100	100	100	100	100	100	100
32	100	100	100	100	100	100	100
33	100	100	100	100	100	100	100
34	100	100	100	100	100	100	100
35	100	100	100	100	100	100	100
36	100	100	100	100	100	100	100

2.J. Metode MFCC pada Derau White (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	90.91	0	0
2	100	100	100	100	100	0	0
3	100	100	100	100	0	0	0
4	100	100	100	100	90.91	0	0
5	100	100	100	100	100	81.82	0
6	100	100	100	100	18.18	0	0
7	100	100	100	100	0	0	0
8	100	100	100	100	100	100	90.91
9	100	100	100	100	100	0	0
10	100	100	100	100	100	100	81.82
11	100	100	100	100	90.91	0	0
12	100	100	100	100	90.91	0	0
13	100	100	100	100	100	0	0
14	100	100	100	100	100	0	0
15	100	100	100	100	100	100	0
16	100	100	100	100	100	0	0
17	100	100	100	100	36.36	0	0
18	100	100	100	100	100	0	0
19	100	100	100	100	100	100	0
20	100	100	100	100	9.091	0	0
21	100	100	100	100	0	0	0
22	100	100	100	100	9.091	0	0
23	100	100	100	100	100	0	0
24	100	100	100	100	100	81.82	0
25	100	100	100	100	0	0	0
26	100	100	100	100	0	0	0
27	100	100	100	100	9.091	0	0
28	100	100	100	100	9.091	0	0
29	100	100	100	100	90.91	0	0
30	100	100	100	100	100	0	0
31	100	100	100	100	72.73	0	0
32	100	100	100	100	100	100	100
33	100	100	100	100	100	0	0
34	100	100	100	100	100	90.91	0
35	100	100	100	100	0	0	0
36	100	100	100	100	100	18.18	0

2.K. Metode MFCC+SS pada Derau F16 (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	100
2	100	100	100	100	100	100	63.64
3	100	100	100	100	100	100	0
4	100	100	100	100	100	100	81.82
5	100	100	100	100	100	100	100
6	100	100	100	100	100	100	0
7	100	100	100	100	100	100	9.091
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100
11	100	100	100	100	100	100	0
12	100	100	100	100	100	100	90.91
13	100	100	100	100	100	100	100
14	100	100	100	100	100	100	81.82
15	100	100	100	100	100	100	100
16	100	100	100	100	100	100	100
17	100	100	100	100	100	100	90.91
18	100	100	100	100	100	90.91	45.45
19	100	100	100	100	100	100	100
20	100	100	100	100	100	100	45.45
21	100	100	100	100	100	100	18.18
22	100	100	100	100	100	100	0
23	100	100	100	100	100	100	72.73
24	100	100	100	100	100	100	90.91
25	100	100	100	100	100	100	18.18
26	100	100	100	100	100	90.91	0
27	100	100	100	100	100	100	0
28	100	100	100	100	100	100	63.64
29	100	100	100	100	100	100	54.55
30	100	100	100	100	100	100	72.73
31	100	100	100	100	100	100	72.73
32	100	100	100	100	100	100	100
33	100	100	100	100	100	100	81.82
34	100	100	100	100	100	100	90.91
35	100	100	100	100	100	100	0
36	100	100	100	100	100	100	81.82

2.L. Metode MFCC+SS pada Derau Hfchannel (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	45.45
2	100	100	100	100	100	90.91	9.091
3	100	100	100	100	100	9.091	0
4	100	100	100	100	90.91	9.091	0
5	100	100	100	100	100	100	9.091
6	100	100	100	100	81.82	0	0
7	100	100	100	100	90.91	0	0
8	100	100	100	100	100	100	100
9	100	100	100	100	100	90.91	0
10	100	100	100	100	100	100	100
11	100	100	100	100	100	54.55	0
12	100	100	100	100	100	72.73	0
13	100	100	100	100	100	90.91	0
14	100	100	100	100	100	72.73	0
15	100	100	100	100	100	100	100
16	100	100	100	100	100	90.91	27.27
17	100	100	100	100	100	45.45	0
18	100	100	100	100	100	27.27	0
19	100	100	100	100	100	100	90.91
20	100	100	100	100	100	45.45	0
21	100	100	100	100	100	0	0
22	100	100	100	100	81.82	0	0
23	100	100	100	100	100	100	18.18
24	100	100	100	100	100	72.73	0
25	100	100	100	100	100	0	0
26	100	100	100	100	81.82	0	0
27	100	100	100	100	100	0	0
28	100	100	100	100	100	45.45	0
29	100	100	100	100	100	45.45	0
30	100	100	100	100	100	100	0
31	100	100	100	100	100	63.64	0
32	100	100	100	100	100	100	90.91
33	100	100	100	100	100	100	0
34	100	100	100	100	100	100	18.18
35	100	100	100	100	100	0	0
36	100	100	100	100	100	100	0

2.M. Metode MFCC+SS pada Derau Pink (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	100
2	100	100	100	100	100	100	90.91
3	100	100	100	100	100	100	27.27
4	100	100	100	100	100	100	72.73
5	100	100	100	100	100	100	100
6	100	100	100	100	100	100	0
7	100	100	100	100	100	100	0
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100
11	100	100	100	100	100	100	9.091
12	100	100	100	100	100	100	54.55
13	100	100	100	100	100	100	54.55
14	100	100	100	100	100	100	72.73
15	100	100	100	100	100	100	100
16	100	100	100	100	100	100	90.91
17	100	100	100	100	100	100	63.64
18	100	100	100	100	100	100	63.64
19	100	100	100	100	100	100	100
20	100	100	100	100	100	100	36.36
21	100	100	100	100	100	100	9.091
22	100	100	100	100	100	100	0
23	100	100	100	100	100	100	100
24	100	100	100	100	100	100	72.73
25	100	100	100	100	100	100	27.27
26	100	100	100	100	100	100	0
27	100	100	100	100	100	100	9.091
28	100	100	100	100	100	100	81.82
29	100	100	100	100	100	100	100
30	100	100	100	100	100	100	100
31	100	100	100	100	100	100	81.82
32	100	100	100	100	100	100	100
33	100	100	100	100	100	100	100
34	100	100	100	100	100	100	90.91
35	100	100	100	100	100	100	9.091
36	100	100	100	100	100	100	100

2.N. Metode MFCC+SS pada Derau Volvo (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	100
2	100	100	100	100	100	100	100
3	100	100	100	100	100	100	100
4	100	100	100	100	100	100	100
5	100	100	100	100	100	100	100
6	100	100	100	100	100	100	100
7	100	100	100	100	100	100	100
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100
11	100	100	100	100	100	100	100
12	100	100	100	100	100	100	100
13	100	100	100	100	100	100	100
14	100	100	100	100	100	100	100
15	100	100	100	100	100	100	100
16	100	100	100	100	100	100	100
17	100	100	100	100	100	100	100
18	100	100	100	100	100	100	100
19	100	100	100	100	100	100	100
20	100	100	100	100	100	100	100
21	100	100	100	100	100	100	100
22	100	100	100	100	100	100	100
23	100	100	100	100	100	100	100
24	100	100	100	100	100	100	100
25	100	100	100	100	100	100	100
26	100	100	100	100	100	100	100
27	100	100	100	100	100	100	100
28	100	100	100	100	100	100	100
29	100	100	100	100	100	100	100
30	100	100	100	100	100	100	100
31	100	100	100	100	100	100	100
32	100	100	100	100	100	100	100
33	100	100	100	100	100	100	100
34	100	100	100	100	100	100	100
35	100	100	100	100	100	100	100
36	100	100	100	100	100	100	100

2.O. Metode MFCC+SS pada Derau White (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	9.091
2	100	100	100	100	100	100	0
3	100	100	100	100	100	100	0
4	100	100	100	100	100	100	0
5	100	100	100	100	100	100	27.27
6	100	100	100	100	100	81.82	0
7	100	100	100	100	100	9.091	0
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	72.73
10	100	100	100	100	100	100	100
11	100	100	100	100	100	54.55	0
12	100	100	100	100	100	100	9.091
13	100	100	100	100	100	81.82	0
14	100	100	100	100	100	90.91	9.091
15	100	100	100	100	100	100	100
16	100	100	100	100	100	100	9.091
17	100	100	100	100	100	81.82	0
18	100	100	100	100	100	81.82	0
19	100	100	100	100	100	100	100
20	100	100	100	100	100	72.73	0
21	100	100	100	100	100	0	0
22	100	100	100	100	100	45.45	0
23	100	100	100	100	100	100	0
24	100	100	100	100	100	90.91	36.36
25	100	100	100	100	100	90.91	0
26	100	100	100	100	100	0	0
27	100	100	100	100	100	45.45	0
28	100	100	100	100	100	100	0
29	100	100	100	100	100	100	0
30	100	100	100	100	100	100	72.73
31	100	100	100	100	100	100	0
32	100	100	100	100	100	100	100
33	100	100	100	100	100	100	27.27
34	100	100	100	100	100	90.91	9.091
35	100	100	100	100	100	45.45	0
36	100	100	100	100	100	100	18.18

2.P. Metode MFCC+wiener pada Derau F16 (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	90.91
2	100	100	100	100	100	100	54.55
3	100	100	100	100	100	90.91	36.36
4	100	100	100	100	100	100	72.73
5	100	100	100	100	100	100	100
6	100	100	100	100	100	100	100
7	100	100	100	100	100	100	100
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100
11	100	100	100	100	100	100	18.18
12	100	100	100	100	100	100	90.91
13	100	100	100	100	100	100	81.82
14	100	100	100	100	100	100	81.82
15	100	100	100	100	100	100	63.64
16	81.82	81.82	81.82	81.82	81.82	63.64	36.36
17	100	100	100	100	100	90.91	90.91
18	100	100	100	100	100	63.64	45.45
19	100	100	100	100	100	100	90.91
20	100	100	100	90.91	90.91	81.82	63.64
21	100	100	100	100	100	100	90.91
22	100	100	100	100	100	81.82	9.091
23	81.82	81.82	72.73	81.82	72.73	63.64	45.45
24	100	100	100	100	100	100	36.36
25	63.64	63.64	54.55	63.64	63.64	54.55	45.45
26	100	100	100	100	100	72.73	27.27
27	100	100	100	100	100	100	36.36
28	100	100	100	100	100	100	90.91
29	90.91	90.91	90.91	90.91	90.91	90.91	54.55
30	100	100	100	100	90.91	90.91	90.91
31	90.91	100	100	81.82	81.82	81.82	54.55
32	100	100	100	100	100	100	36.36
33	54.55	54.55	54.55	54.55	54.55	54.55	54.55
34	81.82	81.82	81.82	81.82	81.82	81.82	72.73
35	100	100	100	100	100	63.64	9.091
36	100	90.91	100	100	100	90.91	81.82

2.Q. Metode MFCC+wiener pada Derau Hfchannel (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	81.82
2	100	100	100	100	100	100	27.27
3	100	100	100	100	90.91	72.73	45.45
4	100	100	100	100	100	100	36.36
5	100	100	100	100	100	100	72.73
6	100	100	100	100	100	100	90.91
7	100	100	100	100	100	90.91	90.91
8	100	100	100	100	100	100	100
9	100	100	100	100	100	90.91	90.91
10	100	100	100	100	100	100	100
11	100	100	100	100	90.91	36.36	9.091
12	100	100	100	100	100	90.91	72.73
13	100	100	100	100	90.91	54.55	0
14	100	100	100	100	100	81.82	36.36
15	100	100	100	100	81.82	90.91	27.27
16	81.82	81.82	81.82	81.82	81.82	45.45	9.091
17	100	100	100	100	100	81.82	90.91
18	100	100	100	90.91	81.82	45.45	9.091
19	100	100	100	100	100	100	81.82
20	100	100	100	100	81.82	72.73	45.45
21	100	100	100	100	100	90.91	90.91
22	100	100	90.91	90.91	72.73	27.27	9.091
23	81.82	81.82	72.73	72.73	63.64	63.64	0
24	100	100	100	100	100	81.82	18.18
25	63.64	72.73	63.64	54.55	45.45	27.27	0
26	100	100	100	100	100	100	18.18
27	100	100	100	100	100	72.73	0
28	100	100	100	100	100	90.91	81.82
29	90.91	90.91	90.91	90.91	90.91	90.91	36.36
30	100	100	100	100	90.91	90.91	63.64
31	90.91	81.82	90.91	90.91	45.45	36.36	63.64
32	100	100	100	100	81.82	54.55	9.091
33	54.55	54.55	54.55	54.55	54.55	36.36	9.091
34	81.82	81.82	81.82	81.82	81.82	63.64	45.45
35	100	100	100	100	72.73	27.27	0
36	100	90.91	100	90.91	100	90.91	63.64

2.R. Metode MFCC+wiener pada Derau Pink (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	90.91
2	100	100	100	100	100	100	18.18
3	100	100	100	100	100	90.91	45.45
4	100	100	100	100	100	100	63.64
5	100	100	100	100	100	100	100
6	100	100	100	100	100	100	100
7	100	100	100	100	100	100	100
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100
11	100	100	100	100	100	81.82	18.18
12	100	100	100	100	100	100	36.36
13	100	100	100	100	100	100	81.82
14	100	100	100	100	100	90.91	27.27
15	100	100	100	100	100	81.82	63.64
16	81.82	81.82	81.82	81.82	81.82	81.82	36.36
17	100	100	100	100	100	90.91	45.45
18	100	100	100	100	100	54.55	9.091
19	100	100	100	100	100	100	100
20	100	100	100	100	81.82	63.64	27.27
21	100	100	100	100	100	100	90.91
22	100	100	100	100	90.91	36.36	0
23	81.82	72.73	81.82	72.73	72.73	63.64	45.45
24	100	100	100	100	100	81.82	0
25	63.64	63.64	63.64	63.64	54.55	54.55	18.18
26	100	100	100	100	100	45.45	0
27	100	100	100	100	100	100	27.27
28	100	100	100	100	100	81.82	9.091
29	90.91	90.91	90.91	90.91	100	72.73	18.18
30	100	90.91	90.91	100	100	81.82	72.73
31	90.91	90.91	81.82	72.73	72.73	63.64	45.45
32	100	100	100	100	100	72.73	0
33	54.55	54.55	54.55	45.45	54.55	45.45	36.36
34	81.82	81.82	81.82	81.82	81.82	81.82	27.27
35	100	100	100	100	90.91	72.73	45.45
36	100	100	90.91	100	90.91	90.91	63.64

2.S. Metode MFCC+wiener pada Derau Volvo (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	100	100
2	100	100	100	100	100	100	100
3	100	100	100	100	100	90.91	100
4	100	100	100	100	100	100	100
5	100	100	100	100	100	100	100
6	100	100	100	100	100	100	100
7	100	100	100	100	100	100	100
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100
11	100	100	100	100	100	100	81.82
12	100	100	100	100	100	100	100
13	100	100	100	100	100	100	100
14	100	100	100	100	100	100	90.91
15	100	100	100	100	100	100	100
16	81.82	81.82	81.82	81.82	81.82	81.82	81.82
17	100	100	100	100	100	100	100
18	100	100	100	100	100	100	90.91
19	100	100	100	100	100	100	100
20	100	100	100	100	100	100	81.82
21	100	100	100	100	100	100	100
22	100	100	100	100	100	100	90.91
23	81.82	81.82	81.82	81.82	72.73	81.82	72.73
24	100	100	100	100	100	100	100
25	63.64	63.64	63.64	81.82	63.64	63.64	54.55
26	100	100	100	100	100	100	100
27	100	100	100	100	100	100	100
28	100	100	100	100	100	100	100
29	90.91	90.91	90.91	90.91	90.91	90.91	90.91
30	100	100	100	100	100	90.91	90.91
31	90.91	100	90.91	90.91	90.91	81.82	72.73
32	100	100	100	100	100	100	100
33	54.55	54.55	45.45	54.55	45.45	54.55	54.55
34	81.82	90.91	81.82	81.82	81.82	81.82	81.82
35	100	100	100	100	100	100	90.91
36	100	100	100	100	100	100	90.91

2.T. Metode MFCC+wiener pada Derau White (dalam satuan %)

Pembicara	Tingkat SNR						
	Bersih	25dB	20dB	15dB	10dB	5dB	0dB
1	100	100	100	100	100	90.91	81.82
2	100	100	100	100	100	72.73	9.091
3	100	100	100	100	90.91	81.82	45.45
4	100	100	100	100	100	100	63.64
5	100	100	100	100	100	100	100
6	100	100	100	100	100	100	100
7	100	100	100	100	100	100	100
8	100	100	100	100	100	100	100
9	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100
11	100	100	100	100	100	63.64	0
12	100	100	100	100	100	90.91	27.27
13	100	100	100	100	100	100	27.27
14	100	100	100	100	100	90.91	0
15	100	100	100	100	90.91	90.91	27.27
16	81.82	81.82	81.82	81.82	72.73	45.45	0
17	100	100	100	100	100	81.82	9.091
18	100	100	100	100	90.91	18.18	0
19	100	100	100	100	100	100	100
20	100	100	100	100	81.82	45.45	9.091
21	100	100	100	100	100	100	100
22	100	100	100	100	81.82	45.45	0
23	81.82	81.82	72.73	72.73	72.73	63.64	9.091
24	100	100	100	100	100	45.45	0
25	63.64	63.64	63.64	63.64	63.64	36.36	18.18
26	100	100	100	100	81.82	27.27	9.091
27	100	100	100	100	100	100	9.091
28	100	100	100	100	100	54.55	9.091
29	90.91	90.91	90.91	90.91	90.91	18.18	9.091
30	100	100	100	100	81.82	81.82	45.45
31	90.91	81.82	81.82	90.91	81.82	54.55	18.18
32	100	100	100	100	100	63.64	0
33	54.55	54.55	54.55	54.55	54.55	45.45	45.45
34	81.82	81.82	81.82	81.82	81.82	54.55	9.091
35	100	100	100	100	100	72.73	27.27
36	100	100	100	90.91	90.91	63.64	27.27

LAMPIRAN 3

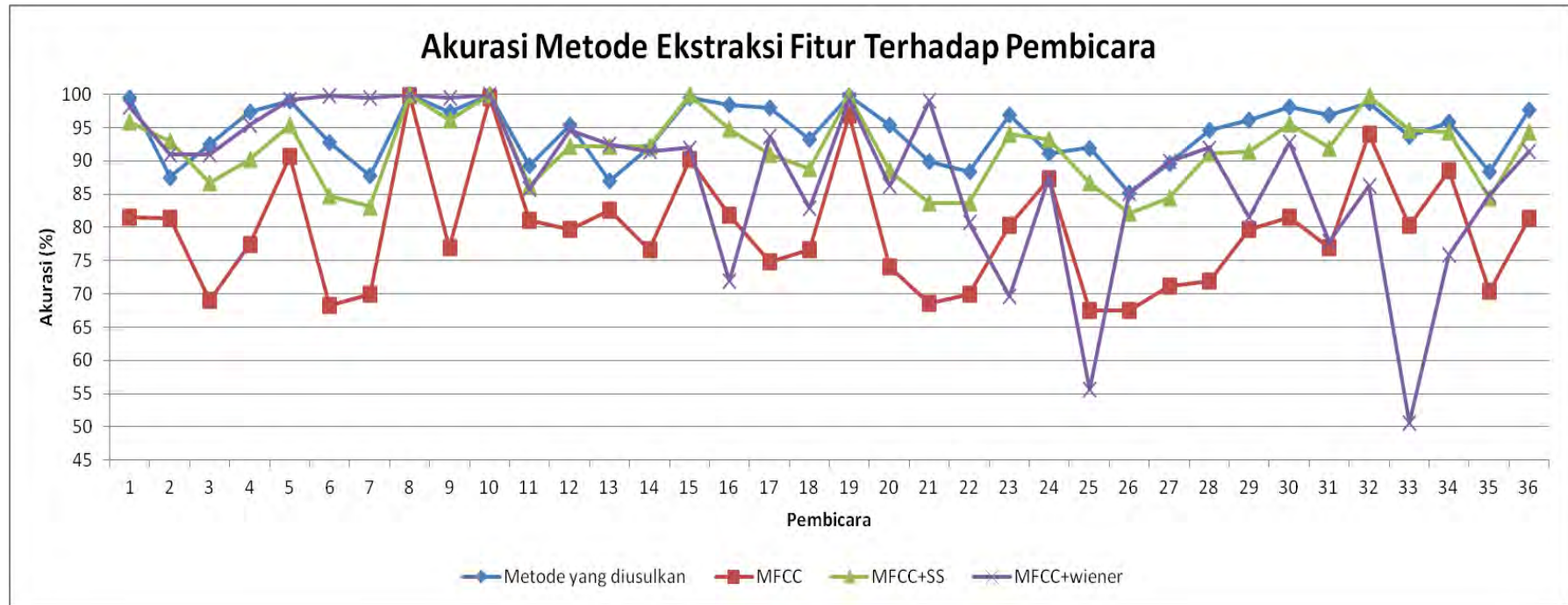
3.A. Tabel Akurasi Metode Ekstraksi Fitur Terhadap Pembicara (dalam satuan %)

Pembicara	Metode yang diusulkan	MFCC	MFCC+SS	MFCC+wiener
1	99.48	81.56	95.84	98.18
2	87.53	81.3	92.99	90.91
3	92.47	69.09	86.75	90.91
4	97.4	77.4	90.13	95.32
5	98.96	90.65	95.32	99.22
6	92.73	68.31	84.68	99.74
7	87.79	69.87	83.12	99.48
8	100	99.74	100	100
9	97.4	76.88	96.1	99.48
10	100	99.48	100	100
11	89.35	81.04	86.23	85.71
12	95.32	79.74	92.21	94.55
13	87.01	82.6	92.21	92.47
14	91.95	76.62	92.21	91.43
15	99.48	90.13	100	91.95
16	98.44	81.82	94.81	71.95
17	97.92	74.81	90.91	93.77
18	93.25	76.62	88.83	82.86
19	99.74	96.88	99.74	99.22
20	95.32	74.03	88.57	86.23
21	89.87	68.57	83.64	98.96
22	88.31	69.87	83.64	80.78
23	96.88	80.26	94.03	69.61
24	91.17	87.27	93.25	87.53
25	91.95	67.53	86.75	55.58
26	85.19	67.53	82.08	85.19
27	89.61	71.17	84.42	89.87
28	94.55	71.95	91.17	91.95
29	96.1	79.74	91.43	81.56
30	98.18	81.56	95.58	92.73
31	96.88	76.88	91.95	77.92
32	98.7	94.03	99.74	86.23
33	93.77	80.26	94.55	50.65

34	95.84	88.57	94.29	75.84
35	88.31	70.39	84.42	84.94
36	97.66	81.3	94.29	91.43

Ket: Nilai akurasi tertinggi

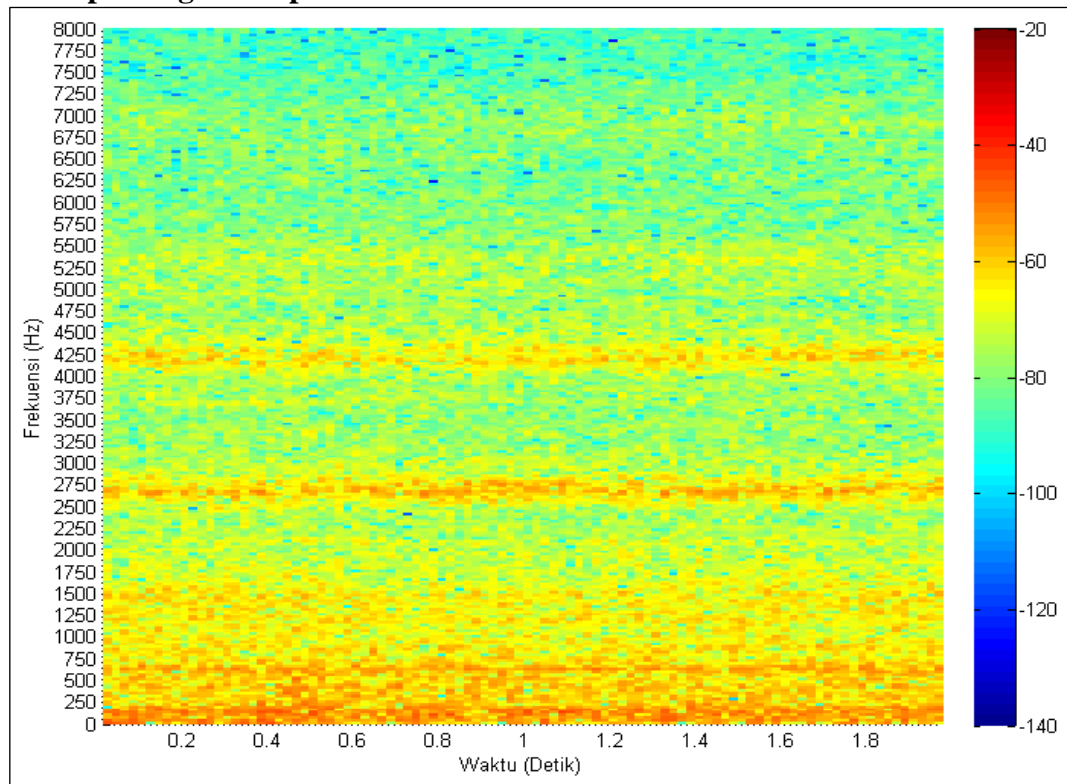
3.B. Grafik Akurasi Metode Ekstraksi Fitur Terhadap Pembicara



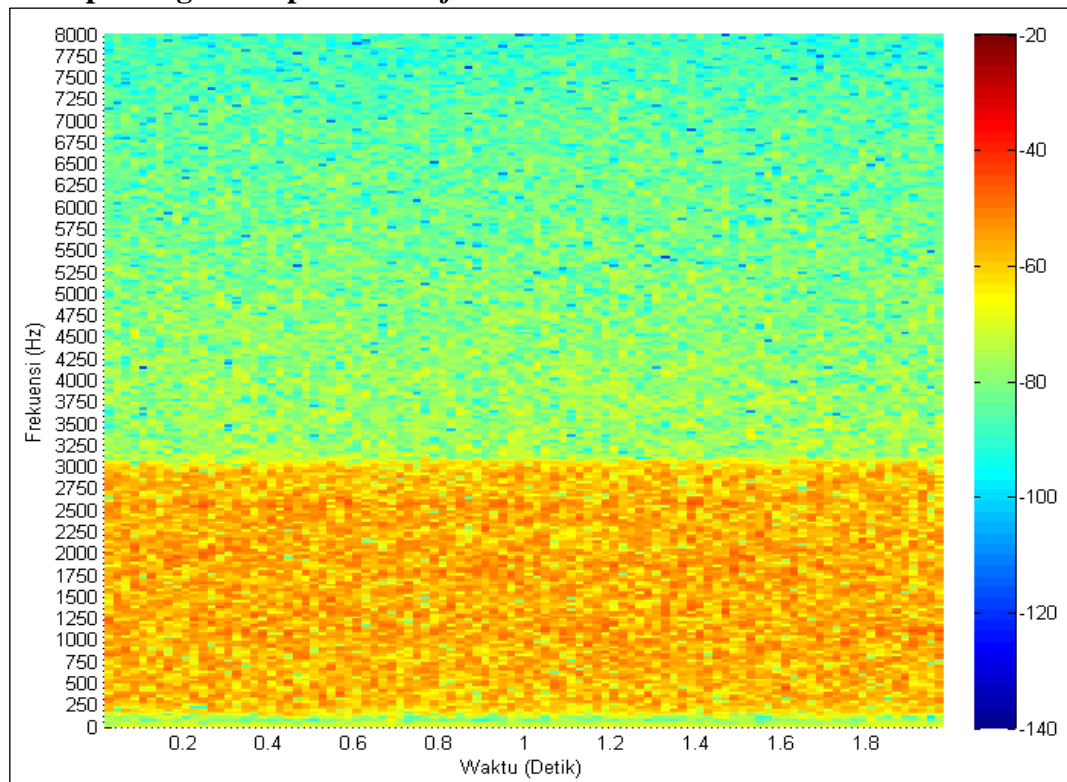
[Halaman ini sengaja dikosongkan]

LAMPIRAN 4

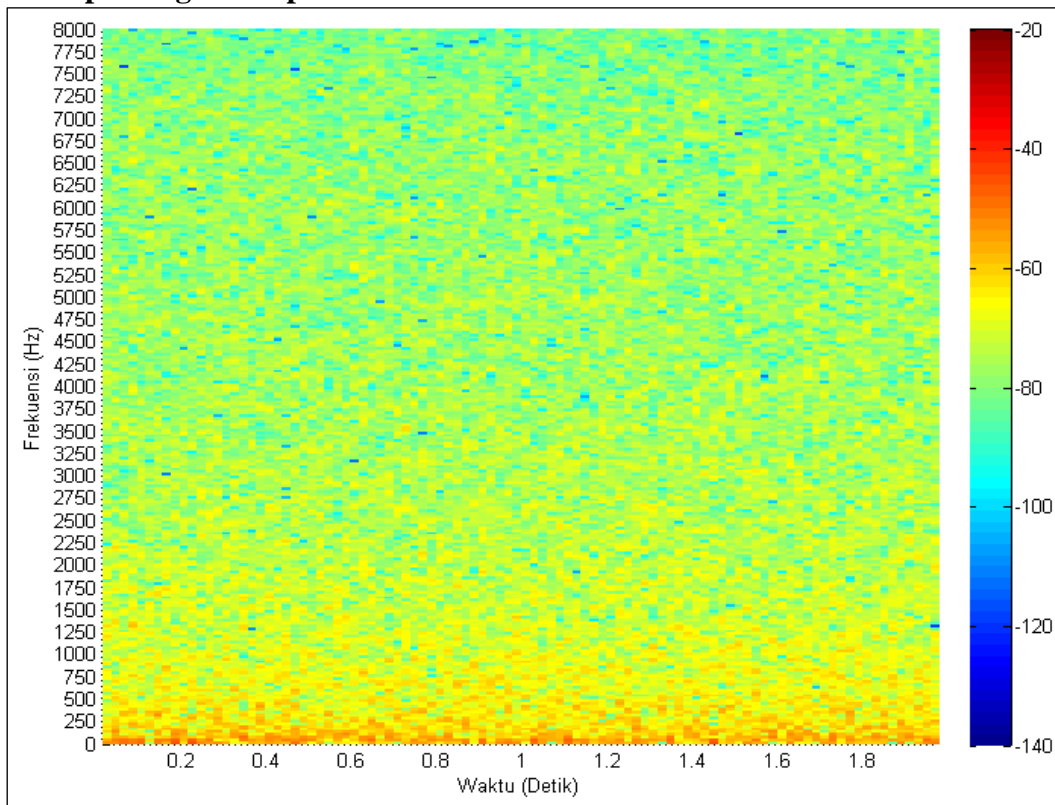
4.A. Spektrogram Tipe Derau *F16*



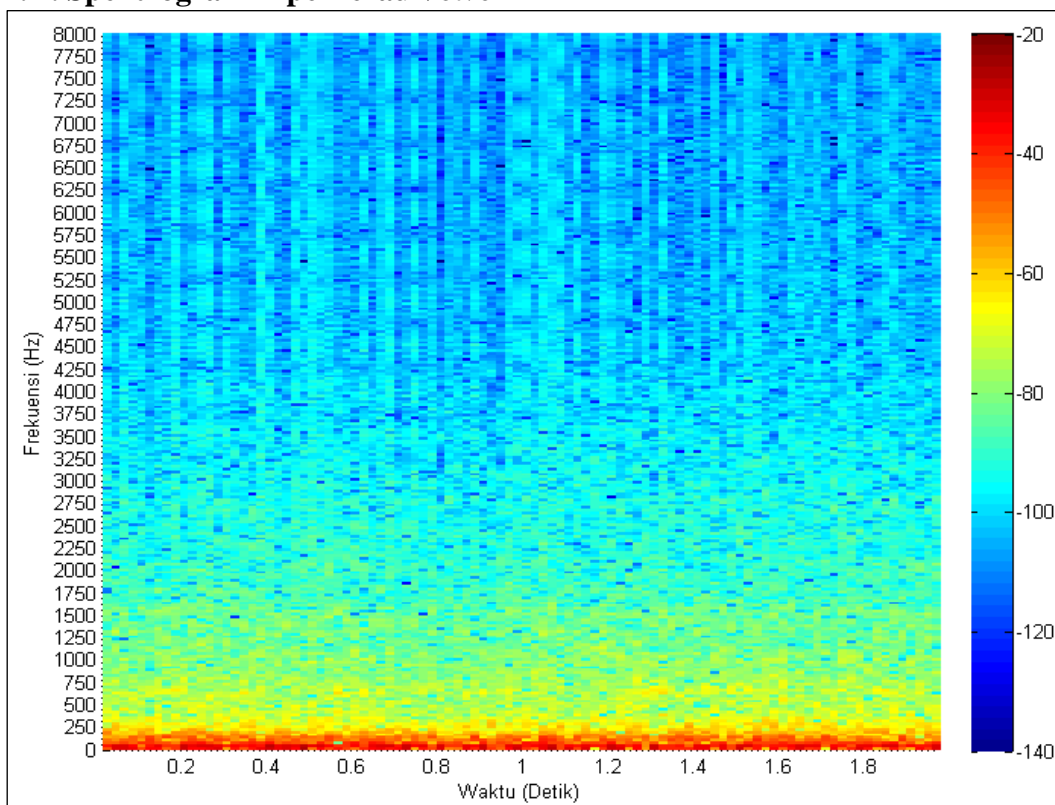
4.B. Spektrogram Tipe Derau *Hfchannel*



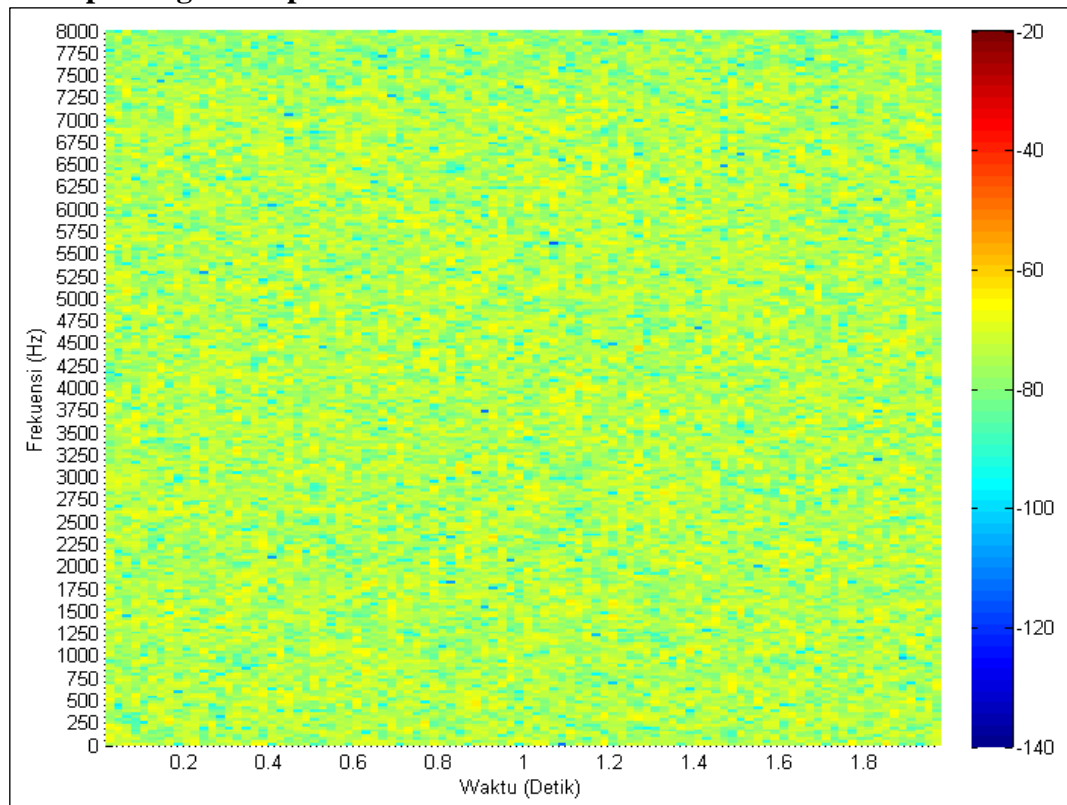
4.C. Spektrogram Tipe Derau *Pink*



4.D. Spektrogram Tipe Derau *Volvo*



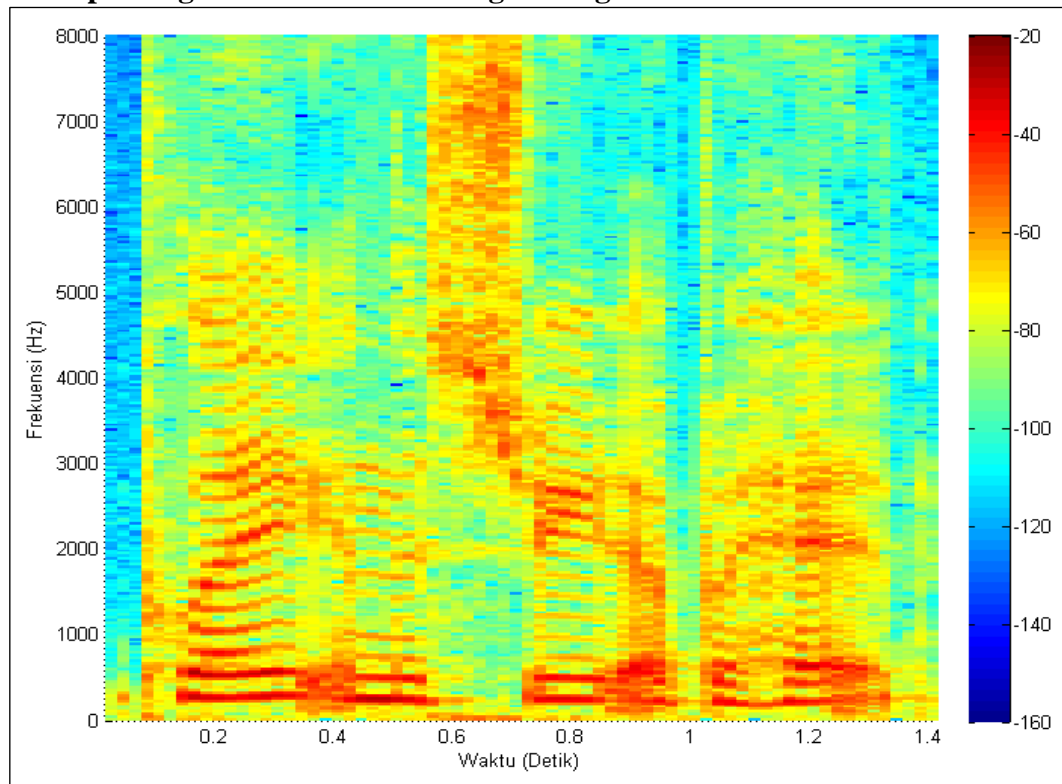
4.E. Spektrogram Tipe Derau *White*



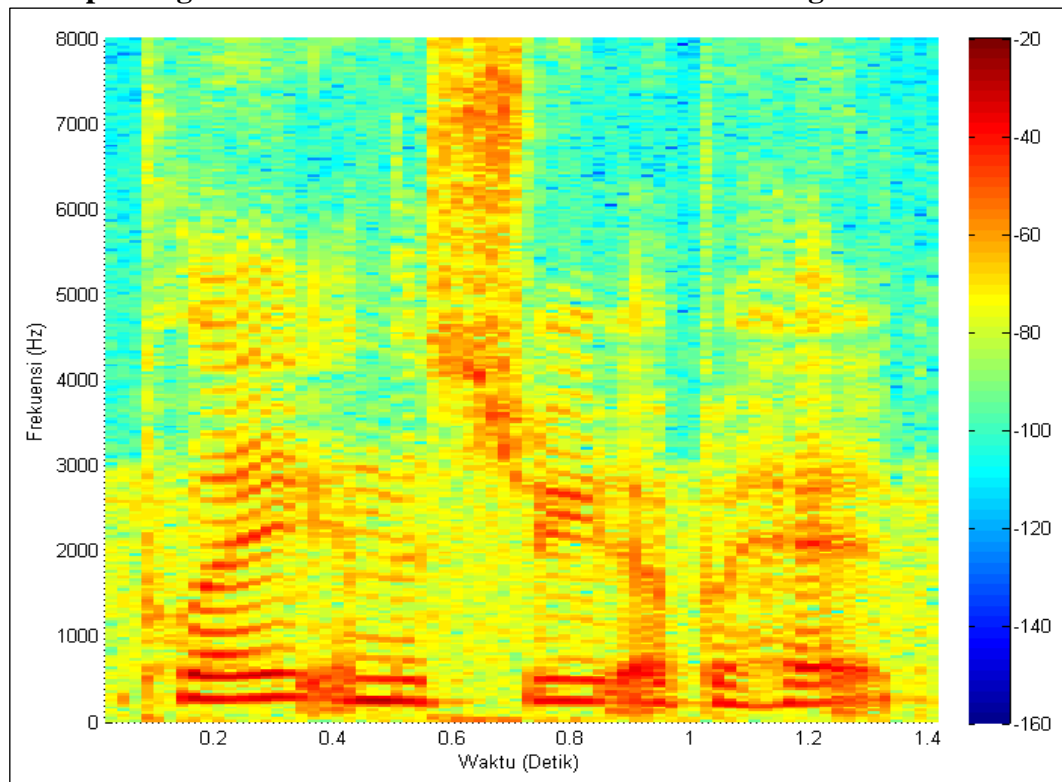
[Halaman ini sengaja dikosongkan]

LAMPIRAN 5

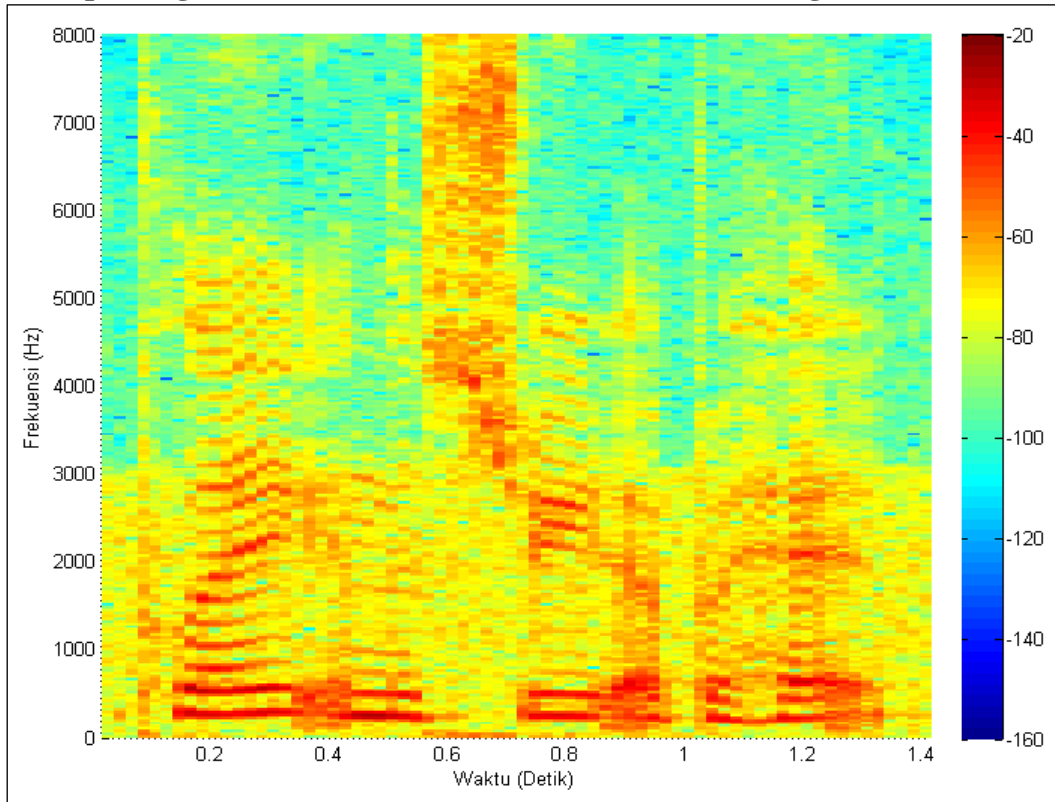
5.A. Spektrogram Pembicara 1 dengan Tingkat SNR Bersih



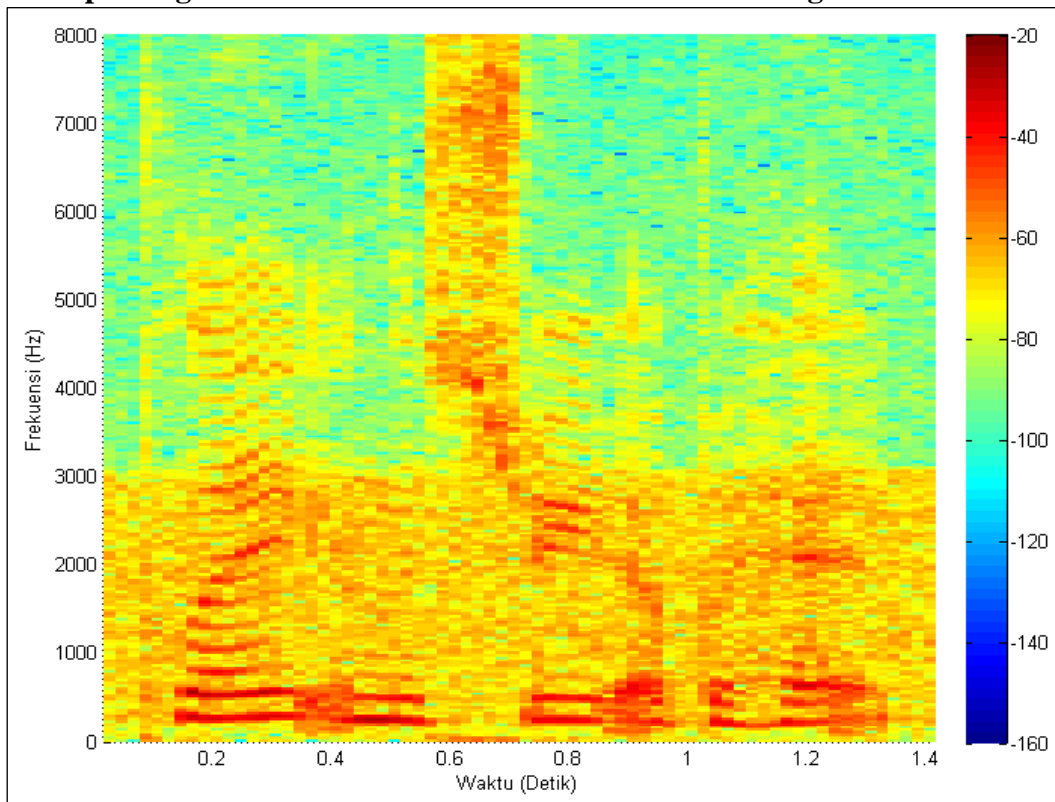
5.B. Spektrogram Pembicara 1 Pada Derau Hfchannel dengan SNR 25dB



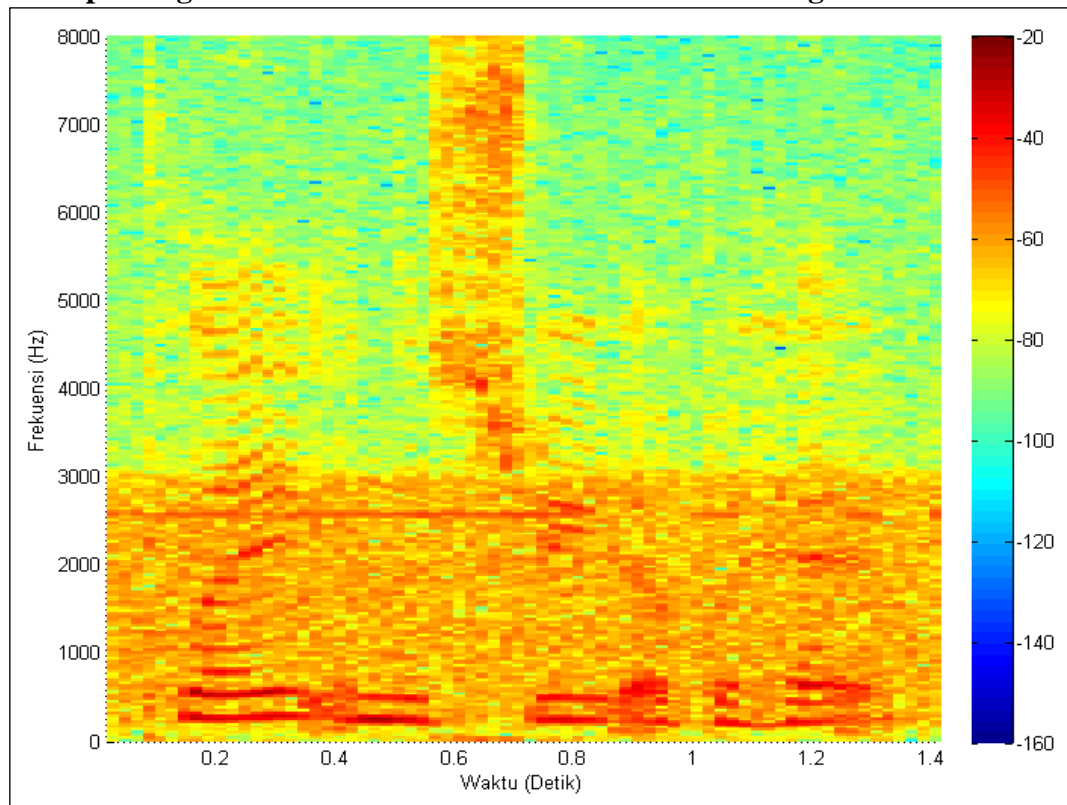
5.C. Spektrogram Pembicara 1 Pada Derau Hfchannel dengan SNR 20dB



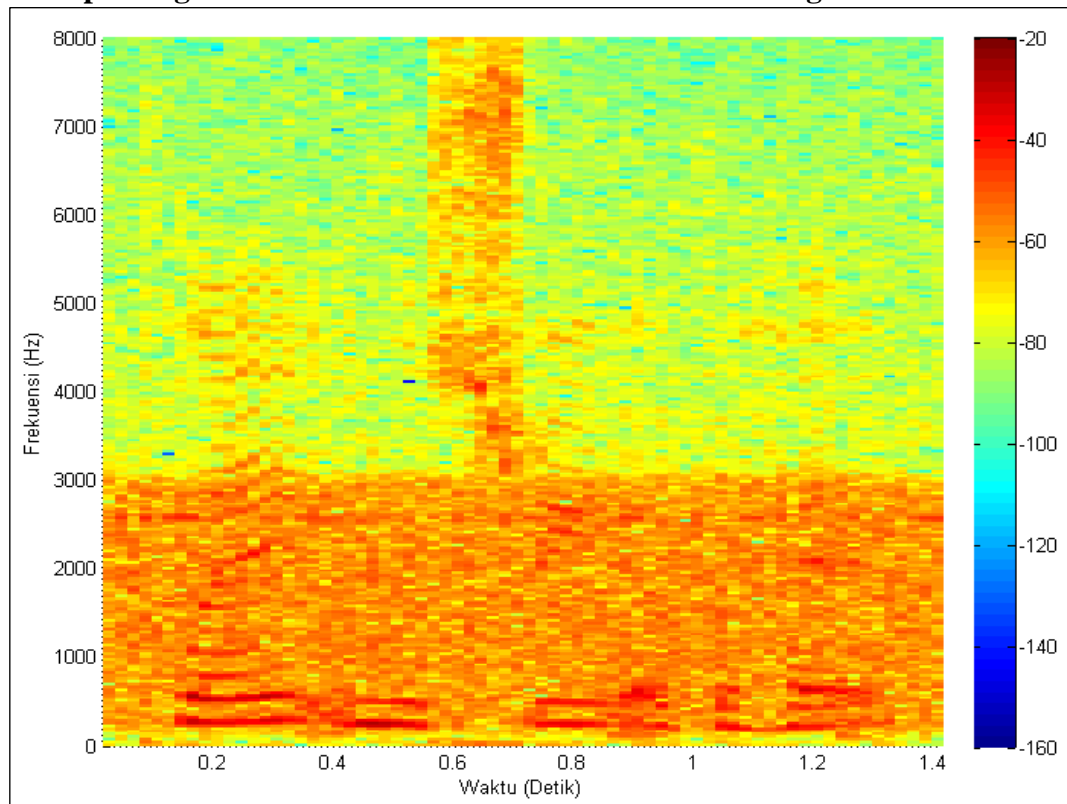
5.D. Spektrogram Pembicara 1 Pada Derau Hfchannel dengan SNR 15dB



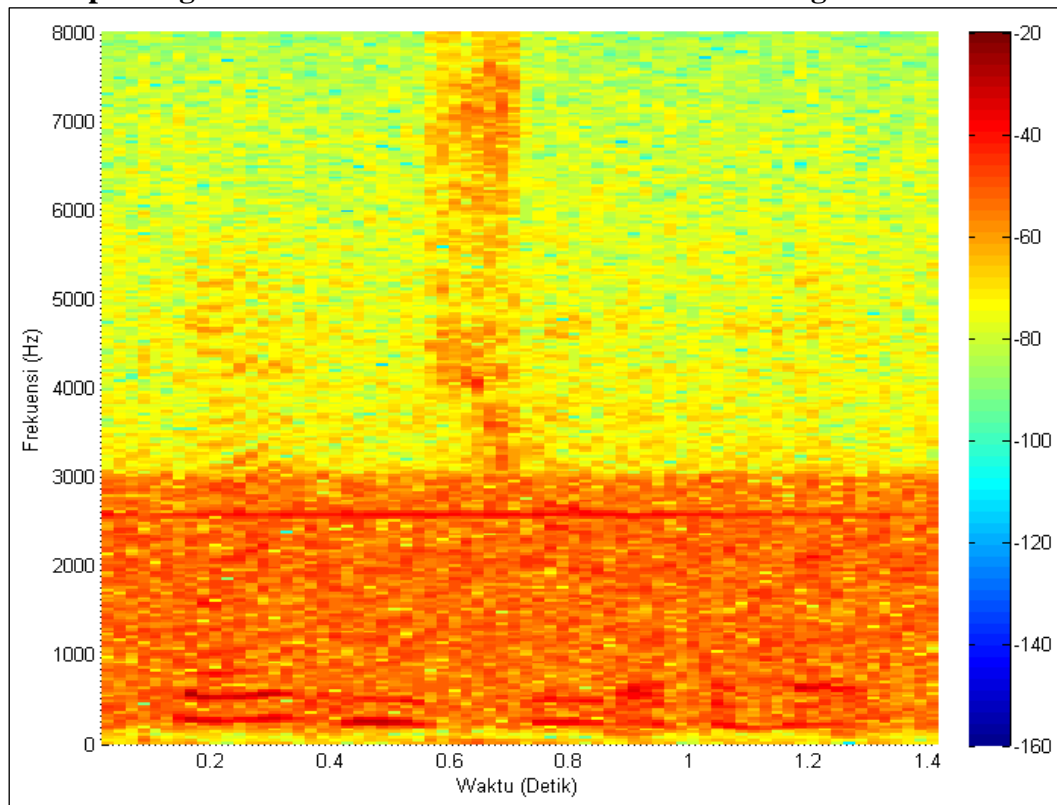
5.E. Spektrogram Pembicara 1 Pada Derau Hfchannel dengan SNR 10dB



5.F. Spektrogram Pembicara 1 Pada Derau Hfchannel dengan SNR 5dB



5.G. Spektrogram Pembicara 1 Pada Derau Hfchannel dengan SNR 0dB



BIODATA PENULIS



Yanuar Risah Prayogi lahir di Sidoarjo pada tanggal 6 Januari 1988. Penulis telah menempuh pendidikan mulai dari SDN 2 Wonoayu (1994-2000), SMPN 1 Wonoayu (2000-2003), SMAN 1 Wonoayu (2003-2006), S1 Universitas Brawijaya (2006-2011), dan pada tahun 2013 penulis meneruskan pendidikan Magister di Teknik Informatika ITS.

Setelah lulus S1 penulis bekerja sebagai *senior software engineer* pada bagian *mobile* dan *server* di perusahaan swasta. Kemudian mendapatkan kesempatan untuk melanjutkan S2 di Teknik Informatika ITS.

Dalam perkuliahan, penulis mengambil bidang minat Komputasi Cerdas dan Visualisasi (KCV) dan tertarik pada hal yang berhubungan dengan kecerdasan buatan dan *machine learning*.